Measurement of photon plus jets production and identification of boosted top quarks in pp collisions at the LHC using the ATLAS detector

Submitted in partial fulfilment of the requirements
for the degree of Doctor in Physics

Héctor de la Torre Pérez
Supervisors: Claudia Glasman and Juan Terrón

November 6, 2015

UNIVERSIDAD AUTONÓMA DE MADRID

Department of Theoretical Physics

You can't take the sky from me

Contents

Introduction

One of the main objectives to build the Large Hadron Collider (LHC) was to search for the Higgs boson, the missing particle of the Standard Model (SM) of elementary particles, in a wide range of masses. The high luminosity and large centre-of-mass energy provided by the proton-proton collisions at the LHC up to 2012 propitiated the discovery of the Higgs boson with a mass of 125 GeV in 2012 by the two main experiments, ATLAS and CMS. The Nobel Prize was awarded in 2013 to François Englert and Peter W. Higgs "for the theoretical discovery of a mechanism that contributes to our understanding of the origin of mass of subatomic particles, and which recently was confirmed through the discovery of the predicted fundamental particle, by the ATLAS and CMS experiments at CERN's Large Hadron Collider".

The LHC is also an ideal environment to test the predictions of the SM, both in the electroweak and strong sectors, and to search for physics beyond the SM (BSM). A large number of measurements with final states involving photons, jets, heavy quarks and W and Z bosons have been performed by both Collaborations to probe the SM and beyond.

In this dissertation, two analyses, involving photons, jets and top quarks and performed using data collected by the ATLAS detector, are described. These analyses provide stringent tests of the SM using prompt photons in association with jets and a novel method to identify and reconstruct top quarks with large transverse momentum (boosted tops).

One of the most important decay channels of the Higgs boson is that into two photons. One of the largest background process for this decay channel is the production of prompt photons in association with jets via QCD. The production of photon plus jets is also a background to searches BSM with photons in the final state. Therefore, measurements of the production of photons plus jets are crucial to constrain the underlying theory. The production of prompt photons in association with jets in pp collisions provides a testing ground for perturbative QCD (pQCD) in a cleaner environment than jet production since the photon originates directly from the hard interaction. The measurements of angular correlations between the photon and the jets can be used to probe the dynamics of the hard-scattering process and colour coherence effects. Since the dominant production mechanism in pp collisions at the LHC proceeds via the $qg \to q\gamma$ process, measurements of prompt-photon plus jet production can be used to constrain the gluon density in the proton, tune the Monte Carlo (MC) models and test the t-channel quark exchange. The distribution of the scattering angle in the photon-jet centre-of-mass ($\theta^{\gamma \mathrm{j}}$) is sensitive to the spin of the exchanged particle and constitutes a fundamental prediction of QCD. Colour connection between partons in the initial and final state modifies the pattern of parton radiation around the final-state partons. This effect, known as colour coherence, has been studied before at Tevatron using dijet events and comparing to predictions with and without such effects. Photon plus jet events are ideal to study this type of effects using only data since the final-state parton can be colour connected while the photon is colourless.

The top quark plays an important role in the SM and in many theories BSM due to its large mass. Top-quark production is also sensitive to the presence of new physics; new massive particles decaying into top quarks with large transverse momentum are possible at the LHC due the large centre-of-mass energy provided in the pp collisions. In the SM, the top quark

decays almost exclusively into a W boson and a b quark. The signature of a $t\bar{t}$ decay is therefore determined by the W boson decay modes. In the semi-leptonic decay mode, one W boson decays into an electron and a neutrino and the other decays hadronically. The hadronically decaying boosted top quark would have very collimated decay products. The identification of such top quarks is difficult by conventional methods. Therefore, new methods based on the reconstruction of a "large-R jet", which clusters all decay products into a single jet, were developed. Substructure techniques to identify and characterise such products are then applied to resolve the three objects inside the jet.

The isolated photon plus jets analysis was performed using the $\sqrt{s} = 8$ TeV 2012 ATLAS dataset. The kinematics of the photon plus one, two and three jets was studied via the measurements of the differential cross sections as functions of the photon transverse energy (E_{T}^{γ}), the p_{T} of the leading, subleading and subsubleading jets. The dynamics of the photon plus one jet system were probed by measuring the cross sections as functions of the photon-jet invariant mass ($m^{\gamma j}$) and $|\cos\theta^{\gamma j}|$. In addition, measurements as functions of the differences between the azimuthal angles of the photon and the jet ($\Delta\phi^{\gamma \mathrm{j}}$) and between the jets ($\Delta\phi^{\mathrm{jj}}$) were also performed. Colour-coherence effects in photon plus two jets events were studied by measuring the angle of the sub-leading jet in the $\eta - \phi$ plane around the photon or the leading jet. The measurements included in this dissertation constitute the first measurement of such effects in ATLAS. The scale evolution of the photon plus one jet system was also studied by measuring the cross sections as functions of $\cos\theta^{\gamma j}$ in different regions of $m^{\gamma j}$. For photon plus two- and three-jet events, the scale evolution was investigated by measuring the angular correlations in different regions of E_{T}^{γ}. Next-to-leading-order (NLO) QCD calculations were compared to the measurements when available.

The identification of top quarks and the reconstruction of its mass via the substructure analysis presented in this dissertation is based on the combined application of the k_{T} and anti-k_{T} jet algorithms and is used in place of other widespread substructure methods such as "trimming" or "pruning" to identify the hard constituents within a "large-R jet". The method developed involves an infrared- and collinear-safe technique to remove the soft contributions from the underlying event and pileup inherent to pp collisions. The method is described in detail and was used to identify top pair production ($t\bar{t}$) with the $\sqrt{s} = 7$ TeV 2011 ATLAS dataset. The kinematics of the reconstructed top quark were also studied as well as the identification efficiency and background rejection factor.

Introducción en castellano

Uno de los principales objetivos para la construcción del gran colisionador hadrónico (LHC, de sus siglas en inglés) fue la búsqueda en un amplio rango de masas del bosón de Higgs: la última partícula necesaria para completar el modelo estándar de las partículas fundamentales. La alta luminosidad instantánea y la alta energía en el centro de masas proporcionadas por las colisiones protón-protón en el LHC propiciaron el descubrimiento del bosón de Higgs en 2012 con una masa de 125 GeV por las colaboraciones ATLAS y CMS. El premio nobel de física fue concedido a François Englert y Peter W. Higgs en 2013 "por el descubrimiento teórico del mecanismo que contribuye a nuestro conocimiento del origen de la masa de las partículas fundamentales y que ha sido confirmado recientemente, mediante el descubrimiento de la partícula fundamental predicha, por los experimentos ATLAS y CMS en el gran colisionador hadrónico en el CERN".

El LHC es el escenario ideal para contrastar las predicciones del modelo estándar, en los sectores fuerte y electrodébil, y para buscar física más allá del modelo estándar. Las colaboraciones ATLAS y CMS han realizado multitud de medidas sobre estados finales constituidos por fotones, jets, quarks pesados y bosones W y Z.

En esta tesis se describen dos análisis sobre fotones, jets y quarks top, realizados usando datos obtenidos con el detector ATLAS. Dichos análisis buscan proporcionar tests rigurosos del modelo estándar usando fotones producidos en asociación con jets (fotón+jets) y desarrollar un método novedoso para identificar y reconstruir quarks top con un momento transverso grande.

Uno de los canales de desintegración más importantes del bosón de Higgs consiste en la desintegración del mismo en dos fotones, siendo una de las fuentes de ruido de fondo más grandes para dicho canal la producción de fotón+jets. Dicha producción es también una de las fuentes de ruido de fondo más importantes para aquellas búsquedas de física más allá del modelo estándar que contengan fotones en el estado final. Por ambas razones, las medidas de la producción de fotón+jets en colisiones protón-protón son cruciales para constreñir las predicciones teóricas. Dicha producción proporciona además tests muy estrictos de cromodinámica cuántica (QCD de sus siglas en inglés) perturbativa en un entorno más limpio que la producción de jets, ya que el fotón proviene directamente de la interacción. También es posible utilizar las medidas de producción de fotón+jets para ajustar los modelos Monte Carlo (MC), estudiar la dinámica del sistema formado por el fotón y el jet, comprobar el intercambio de quarks en el canal t y constreñir la densidad gluónica en el fotón. Este último estudio es posible gracias a que el mecanismo de producción dominante en colisiones protón-protón en el LHC es el proceso $qg \to q\gamma$.

La distribución del ángulo de dispersión en el centro de masas del fotón y el jet ($\theta^{\gamma \mathrm{j}}$) es sensible al espín de la partícula intercambiada y constituye una predicción fundamental de QCD. El flujo de color entre partones en el estado inicial y final modifica el patrón de radiación partónica alrededor de los partones en el estado final. Este fenómeno, conocido como coherencia de color, ha sido estudiado con anterioridad en el acelerador Tevatron usando sucesos con dos jets y comparando los resultados con predicciones teóricas realizadas con y sin dichos efectos. El estudio de sucesos de producción de fotón+jets permite estudiar este tipo de fenómeno usando sólo datos, ya que, en el estado final, el partón puede interaccionar con el flujo de color mientras

que el fotón es incoloro.

El quark top juega un papel importante en el modelo estándar y en muchas teorías más allá del modelo estándar por su gran masa. De igual modo, la producción de quarks top es sensible a la presencia de nueva física; nuevas partículas masivas desintegrándose en quarks top con un gran momento transverso son posibles en el LHC gracias a la alta energía en el centro de masas proporcionada por las colisiones protón-protón. En el modelo estándar, el quark top se desintegra de forma prácticamente exclusiva en un bosón W y un quark b; por ello, la señal de la desintegración de una pareja de quarks top está determinada por los modos de desintegración del bosón W. En el canal de desintegración semi-leptónico, uno de los bosones W se desintegra en un electrón y un neutrino y el otro se desintegra de forma hadrónica. Un quark top con alto momento transverso que se desintegre hadrónicamente tenderá a tener unos productos de desintegración muy colimados y la identificación de dichos productos por métodos convencionales se vuelve difícil. Para afrontar dicho problema se han desarrollado nuevos métodos basados en la reconstrucción de un jet con un radio grande, que junta todos los productos de desintegración en un solo jet. Posteriormente se usan técnicas de sub-estructura para identificar y caracterizar los productos de desintegración, diferenciando los tres objetos que conforman el jet.

El análisis de la producción de fotones en asociación con jets se ha realizado usando los datos obtenidos durante 2012 por el detector ATLAS a una energía en el centro de masas de 8 TeV. La cinemática del fotón en asociación con uno, dos y tres jets se ha estudiado mediante la medida de las secciones eficaces diferenciales como función de la energía transversa del fotón (E_{T}^{γ}) y el momento transverso (p_{T}) del jet con más p_{T}, del segundo jet con más p_{T} y del tercer jet con más p_{T}. La dinámica del fotón en asociación con un jet se ha estudiado mediante la medida de la sección eficaz diferencial como función de la masa invariante entre el fotón y el jet $m^{\gamma j}$ y de $\cos\theta^{\gamma j}$. También se han realizado medidas como función de las diferencias en el ángulo acimutal entre el fotón y los jets ($\Delta\phi^{\gamma \mathrm{j}}$) y de los jets entre sí ($\Delta\phi^{\mathrm{jj}}$).

Se ha estudiado la coherencia de color mediante la medida del ángulo en el plano $\eta - \phi$ entre el segundo jet con más p_{T} y el fotón o entre el segundo jet con más p_{T} y el jet con más p_{T}. Dichos estudios constituyen la primera medida de este tipo de efectos en ATLAS.

Se ha estudiado la evolución con la escala de energía del sistema foton+jet mediante la medida de la sección eficaz diferencial como función de $\cos\theta^{\gamma j}$ para distintas regiones de $m^{\gamma j}$. Para sucesos con dos y tres jets, la evolución con la escala de energía se ha estudiado mediante la medida de las distribuciones angulares en distintas regiones de E_{T}^{γ}. Cálculos teóricos de cromodinámica cuántica al siguiente orden en teoría de perturbaciones han sido comparados con las medidas en aquellos casos en los que han estado disponibles.

Las técnicas de subestructura que se describen en esta tesis para la identificación de quarks top de alto momento transverso y la reconstrucción de su masa están basadas en la combinación de los algoritmos de jet k_{T} y anti-k_{T} y se usan en lugar de otras técnicas de subestructura comúnmente utilizadas como "trimming" o "pruning" para identificar los productos de desintegración dentro de un jet con radio grande. El método desarrollado elimina las contribuciones suaves debidas al suceso subyacente y otras colisiones simultaneas, siempre presentes en las colisiones proton-proton, usando una técnica que no tiene singularidades en los límites infrarrojo y colineal. Dicho método, descrito en detalle en esta tesis, se ha utilizado para la identificación de sucesos de producción de parejas de quarks top, usando para ello los datos obtenidos durante 2011 por el detector ATLAS a una energía en el centro de masas de 7 TeV.

Chapter 1

Theoretical framework

In this chapter, the theoretical framework relevant for the two analyses included in this dissertation, the high transverse-momentum top quark analysis and the isolated-photon plus jets analysis, are introduced.

1.1 The Standard Model

The Standard model (SM) [1–3] of elementary particle physics provides a fundamental description of all elementary particles, their dynamics and interactions. Fermions, particles of half-integer spin, form all matter in nature while the interactions between fermions are interpreted using the exchange of force-mediating bosons of integer spin. Figure 1.1 shows all the particles that are included in the SM and their most important properties. For each particle exist an antiparticle with the same mass but opposite intrinsic quantum numbers.

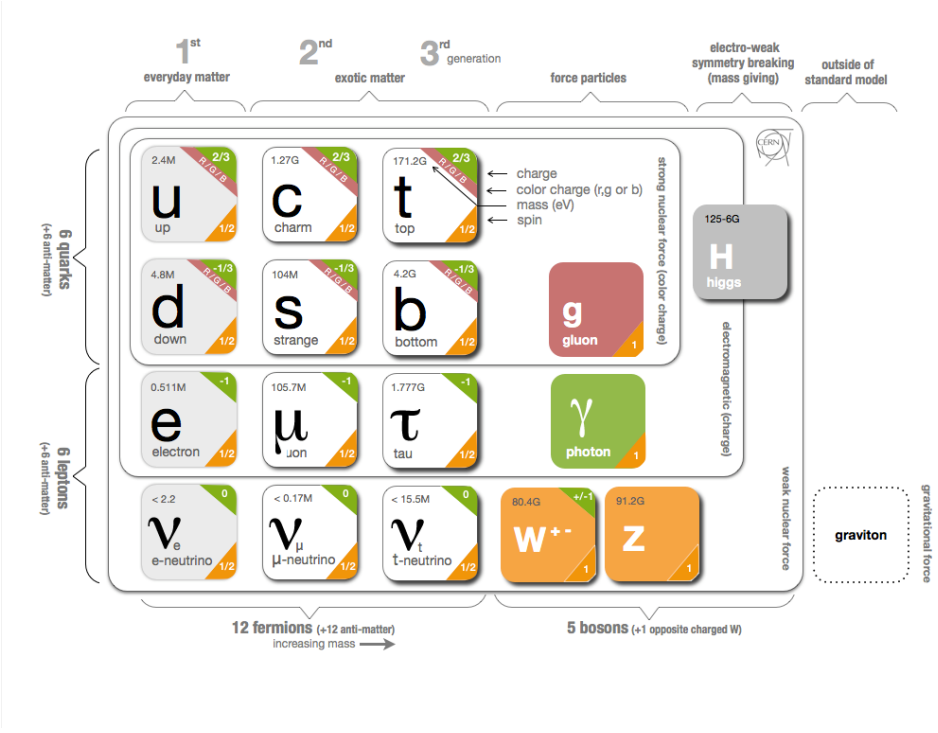


Figure 1.1: *Particles of the Standard Model.*

In principle, the particles embedded into the SM are massless, unlike the observed particles.

Thus, a mechanism must be introduced in the theory to give the particles mass. The W and Z bosons acquire mass through a mechanism called electroweak symmetry breaking (EWSB). In the SM, the Higgs scalar field induces a spontaneous breaking of the electroweak symmetry (EW) when it acquires a non-vanishing vacuum expectation value. The mechanism is commonly referred to as the Brout-Englert-Higgs mechanism [4–7]. The Higgs field also provides mass to the fermions of the SM via Yukawa interactions. The Brout-Englert-Higgs mechanism postulates the existence of one scalar particle, the Higgs boson, which was experimentally discovered during 2012 by the ATLAS and CMS Collaborations at the Large Hadron Collider (LHC) with a mass around 125 GeV [8, 9]. After the discovery, François Englert and Peter Higgs were awarded the Nobel prize in physics *for the theoretical discovery of a mechanism that contributes to our understanding of the origin of mass of subatomic particles, and which recently was confirmed through the discovery of the predicted fundamental particle, by the ATLAS and CMS experiments at CERN's Large Hadron Collider* [10].

The SM is a non-abelian gauge theory, described by the gauge group:

$$SU(3)_c \otimes SU(2)_L \otimes U(1)_Y \tag{1.1}$$

The local $SU(2)_L \otimes U(1)_Y$ symmetry corresponds to the EW interaction, described by the Glashow-Salam-Weinberg model (GSW model), combining the electromagnetic and weak interactions. The remaining $SU(3)_C$ reflects the symmetry of the strong interaction, also known as Quantum Chromodynamics (QCD), namely an exact symmetry of three colours of the six quarks [11–13].

1.2 Physics at the LHC

One of the most crucial elements of any physics analysis at the LHC is the fact that the colliding particles, the protons, are not elementary. The proton is a baryon, a type of composite particle which, according to QCD, is composed of three quarks (or antiquarks). These are called valence quarks and determine the charge and flavour of the baryon. In the case of the proton, the three valence quarks are *uud*, that is, two quarks of type up and one quark of type down.

However, the complete picture of the proton is more complex. Due to the QCD interactions between quarks, gluons are radiated by the valence quarks and split into quark-antiquark pairs, which makes it possible to find any quark flavour inside the proton. This makes the proton behave like a complex system formed by partons (gluons and quarks).

Using the factorisation theorem [14], which expresses the independence of soft structure on the nature of the hard process, originally formulated for Deep Inelastic Scattering (DIS), it is possible to express the experimentally observable cross-section for hadron-hadron scattering at a centre of mass energy $\sqrt{s}$ as a convolution of a non-perturbative but universal (process independent) parton density function (PDF) and a perturbatively calculable partonic scattering cross section:

$$\sigma_{pp\to X}(s) = \sum_{i,j} \int_0^1 dx_1 f_{i/p}(x_1, \mu_F^2) \int_0^1 dx_2 f_{j/p}(x_2, \mu_F^2) \sigma_{ij\to X}(x_1, x_2, s, \mu_F^2) \tag{1.2}$$

where $f_{i/p}(x_i)$ is the proton PDF for the parton species i as a function of the momentum fraction x_i (with respect to the hadron), i and j run over the combination of partons capable of taking part in the process $ij \to X$ and $\sigma_{ij\to X}(x_1, x_2, s, \mu_F^2)$ is the partonic cross section, that knows nothing of the target hadron apart from the fact that it contained the struck parton.

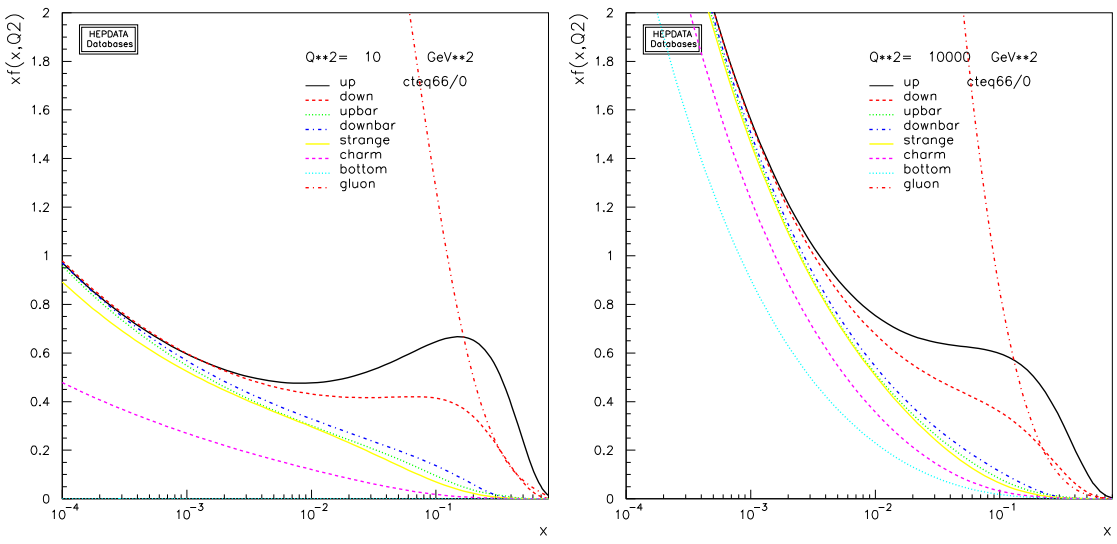
The dividing line between the two is drawn at an arbitrary scale μ_F, called the factorisation scale. Although the factorisation theorem is widely used in many hadron-hadron processes, a

mathematical proof of its validity only applies to inclusive cross-sections in DIS and in Drell-Yan processes [15]. For a general process in hadron-hadron collisions, factorisation is only an ansatz.

1.2.1 Parton densities

The parton density function $f_{i/h}(x_i, \mu_F^2)$ represents the effective density of partons of type i, as a function of the momentum fraction x_i when a hadron of type h is probed at the factorisation scale μ_F. The PDFs are non-perturbative functions which are not a priori calculable and have to be determined experimentally. However, perturbative equations governing their evolution with μ_F can be obtained by requiring that experimental scattering cross sections, such as the one in equation 1.2, be independent of μ_F. The resulting renormalisation group equation (RGE) obtained from this requirement is called the DGLAP (Dokshitzer–Gribov–Lipatov–Altarelli–Parisi) equation [16–18] and can be used to evolve the PDFs from one scale to another. Therefore it is only needed to determine the form of the PDF as a function of x at a single scale μ_0 and then use the DGLAP equations to obtained them at any scale.

There are several collaborations dedicated to obtain PDFs parameterisations (such as CTEQ, MSTW, NNPDF) that use different data and methods to perform fits. One of such sets, at two different scales, for several different parton types is shown in figure 1.2. The figure was obtained using the online HepData tool [19].



(a) (b)

Figure 1.2: *Proton PDFs as functions of the momentum fraction for several types of partons at a factorisation scale $\mu_F^2 = Q^2 = 10$ GeV^2 (a) and $\mu_F^2 = Q^2 = 10^4$ GeV^2 (b). The PDFs were obtained with the HEPDATA online tool and they are part of the CTEQ6.6 PDF set [20].*

1.2.2 Fixed order QCD calculations

In general, a partonic-level cross section can be calculated using perturbative methods and the corresponding Feynmam diagrams and rules of the desired theory. In the case of QCD, the cross-section can be expressed as:

$$\sigma = \tilde{\sigma}^0 + \alpha_s \cdot \tilde{\sigma}^1 + \alpha_s^2 \cdot \tilde{\sigma}^2 + O(\alpha_s^3) \tag{1.3}$$

where $\tilde{\sigma}^0$ is referred to as the leading order (LO) contribution, $\tilde{\sigma}^1$ the next-to-leading order (NLO) contribution and so on. α_s is the coupling of the theory which needs to be small in order

for perturbation theory to work.

When calculating the terms of equation 1.3 using directly the appropriate Feynman rules, several sources of divergences appear, that is, the calculation is not finite. Some of these divergences cancel among themselves or are absorbed in the PDFs; however, the divergences arising from high momenta in loop corrections, known as ultraviolet (UV) divergences, require special treatment.

To get rid of the UV divergences a method, which is known as renormalisation, is used. There are several schemes of renormalisation, such as the minimal subtraction scheme (MS) or the on-shell scheme. By this method, the divergences are cancelled at any fixed order of perturbative theory, but at the cost of introducing an extra scale parameter μ_R, which can be interpreted as the cut-off scale that separates the divergent behaviour from the now finite perturbation calculation. The renormalisation scale, μ_R, is arbitrary. The most important consequence of renormalisation is that several parameters of the theory, in particular α_s, gain a dependence on μ_R.

Let us consider a dimensionless experimental observable, function of a single scale Q and the coupling constant α_s. After renormalisation, since there is a second scale, the dimensionless variable must now depend on the ratio Q/μ_R. However, since the scale μ_R is arbitrary, the dependence on μ_R must cancel on any experimental observable up to the order to which the calculation is performed. This fact allows the construction of a RGE that dictates the evolution of α_s with the scale Q. Such evolution can be expressed as:

$$\mu_R^2 \frac{\partial \alpha_s(Q^2/\mu_R^2)}{\partial \mu_R^2} = \beta(\alpha_s(Q^2/\mu_R^2)) \tag{1.4}$$

where $\beta(\alpha_s(Q^2/\mu_R^2))$ is known as the β-function. The β-function can be calculated in perturbative QCD and can be written as:

$$\beta(\alpha_s) = -b_0 \cdot \alpha_s^2 + O(\alpha_s^3), \qquad b_0 = \frac{33 - 2n_f}{12\pi}, \tag{1.5}$$

where n_f is the number of active light flavours.

Equation 1.4 can be solved to obtain the explicit evolution of α_s within the approximation used in equation 1.5:

$$\alpha_s(Q^2) = \alpha_s(\mu_R^2) \frac{1}{1 + b_0 \alpha_s(\mu_R^2) \log \frac{Q^2}{\mu_R^2}} \tag{1.6}$$

The negative overall sign of the β function leads to the famous result that the QCD coupling effectively decreases with energy, called asymptotic freedom, which, among other consequences, makes quarks behave as free particles at sufficiently high scales or short distances. The converse is also true for low scales, the numerical value of α_s increases towards smaller energies, causing quarks and gluons to form hadronic colour singlets, which is called colour confinement. The evolution of α_s has been measured by various experiments and is shown in figure 1.3. The world average of α_s at a scale equal to the mass of the Z boson ($\alpha_s(M_z)$) is $\alpha_s(M_z) = 0.1185 \pm 0.006$ [21]. After the experimental confirmation of the evolution of α_s in 2004, David J. Gross, Hugh David Politzer and Frank Wilczek were awarded the Nobel prize *for the discovery of asymptotic freedom in the theory of strong interaction* [22].

Finally, equation 1.6 can also be rewritten in the following form:

$$\alpha_s(Q^2) = \frac{1}{b_0 \log \frac{Q^2}{\Lambda^2}}, \tag{1.7}$$

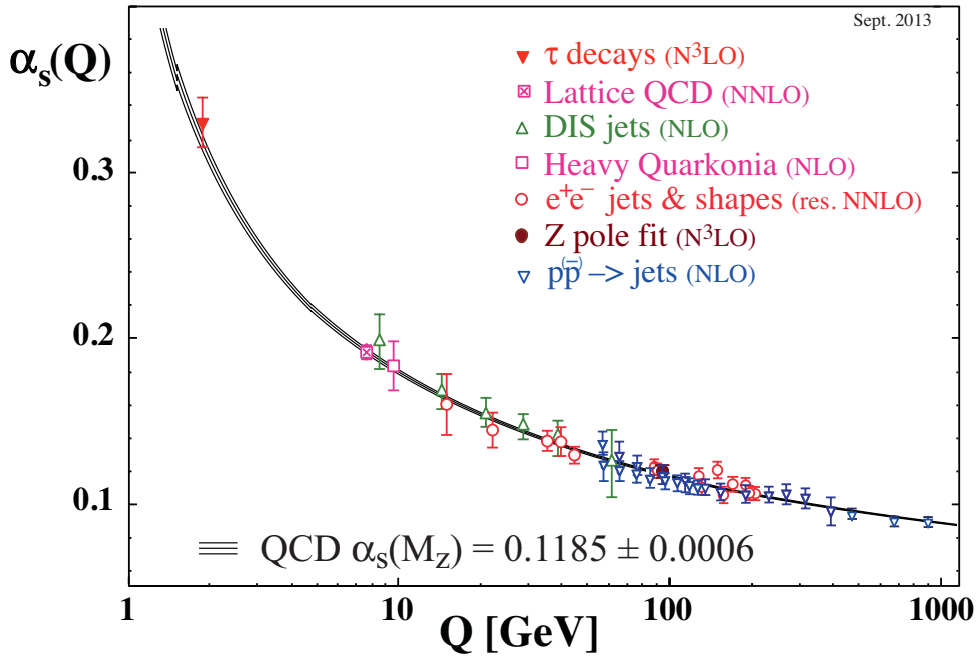


Figure 1.3: *Summary of measurements of α_s as a function of the energy scale Q [21].*

where Λ represents the energy scale at which the perturbative coupling would nominally become infinite and is called the Landau pole. The numerical value of Λ depends on the renormalisation scheme used and takes a value $\sim$ 200 MeV for the modified minimal substraction scheme ($\overline{\text{MS}}$). It is important to note that equation 1.7 only characterizes the purely perturbative results, which are only reliable for $\alpha_s << 1$. In a certain sense, Λ defines the separation between the perturbative region, with free partons, and the non-perturbative region, with strongly coupled partons.

1.3 Jet algorithms

Both of the analyses included in this dissertation, which will be introduced in the following sections, have partons present in the final-state, and thus an experimental tool for handling them is needed. This task is accomplished by jet algorithms.

QCD calculations are performed using perturbation theory for partonic cross sections at high-energy transfers. However, quarks and gluons are not observable, and at large distances these produced partons are confined by the colour force field and forced to dress themselves up into colourless hadrons through a series of processes known as fragmentation and hadronisation, which are certainly non perturbative and will be approached within the context of Monte Carlo simulations (see chapter 3).

The observable counterparts of the partons are jets, collimated bunches of high-energy hadrons which are the result of the fragmentation and hadronisation of the outgoing partons. Currently limited understanding of non perturbative QCD is such that it is not possible to predict the exact patterns of hadrons produced.

While the intuitive definition of a jet as a *large amount of hadronic energy in a small angular region* is sufficient to account for many qualitative features of parton production, any detailed quantitative analysis requires a precise jet definition. A jet algorithm must be able to specify a jet configuration unambiguously, both in theoretical calculations and in experimental practice,

starting from particles detected in the final-state.

In the case of hadron-hadron collisions there is a series of requirements that are necessary to define such a jet algorithm:

1. Simple to use in experimental analyses;
2. Simple to use in theoretical calculations;
3. Infrared and collinear safe;
4. Subject to small hadronisation corrections;
5. Amenable to factorisation of the initial-state collinear singularities into the universal proton PDFs;
6. Not strongly affected by contamination from the beam remnants and the underlying event;
7. Longitudinally invariant.

Requirements 1 and 2 are self-evident. The others follow the motivation of comparing data and theory. In e^+e^- annihilation into hadrons this can be achieved by imposing only the third requirement, i.e. jet cross sections at parton level must be finite order by order in perturbation theory in the limit of massless final-state partons. In the case of initial-state hadrons this is not enough because the center-of-mass energy no longer controls the hardness of the process and, therefore, the jet definition has also to fulfil requirements 4 to 7.

For the analyses presented in this dissertation, two jet algorithms were used, the k_{T} [23] and the anti-k_{T} [24] algorithms, which are described in some detail in the following sections.

1.3.1 The k_{T} algorithm

The inclusive variant of the k_{T} algorithm, in the longitudinally invariant formulation suitable for hadron colliders, works as follows.

In collisions in the hadron-hadron center-of-mass frame with the z-axis taken in the beam direction, the final state of the collision is represented as consisting of a starting set of *protojets* with momenta p_i^μ (e.g. the final-state particles). Starting with the initial list of protojets, the jet algorithm recursively groups pairs of protojets together to form new protojets. The idea is that protojets with nearly parallel momenta should be joined, so that they will eventually form part of the same jet. The algorithm also determines when, for a particular protojet, joining should cease. This protojet is then labelled as a completed "jet" and is not manipulated further. An step-by-step description is listed below:

1. For each pair of protojets i and j, calculate the k_{T} distance, defined as

$$d_{ij} = min(p_{Ti}^2, p_{Tj}^2)\frac{\Delta R_{ij}^2}{R^2} \tag{1.8}$$

$$\Delta R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2 \tag{1.9}$$

   with p_{Ti}, η_i and ϕ_i the transverse momentum, pseudorapidity and azimuth of protojet i, respectively, and R is a jet-radius parameter usually taken of order 1. For each protojet i also work out the distance to the beam $d_{iB} = p_{Ti}^2$.

2. Find the minimum d_{min} of all the values d_{ij},d_{iB}. If d_{min} is a d_{ij} merge protojets i and j into a single protojet, e.g. simply by summing their four momenta. If it is a d_{iB} then declare the protojet i to be a final jet and remove it from the list;

3. Repeat from step 1 till no protojets are left in the list.

The exclusive variant of the k_{T} algorithm is similar, except that:

- when $d_{i,B}$ is the smallest value, the particle is considered to become part of the beam jet (i.e. is discarded);
- clustering is stopped when all d_{ij} and d_{iB} are above some d_{cut} or when some number of protojets is reached.

Reclustering an already reconstructed jet up to a certain d_{cut} or down to a certain number of protojets effectively breaks a jet into smaller jets, which are known as "subjets".

Unless otherwise indicated, the k_{T} algorithm is always used in its inclusive variant in this dissertation.

1.3.2 The anti-k_{T} algorithm

The anti-k_{T} algorithm works in a similar way to the inclusive version of the k_{T} algorithm, with a crucial difference, the d_{ij} distance is calculated as

$$d_{ij} = min(1/p_{Ti}^2, 1/p_{Tj}^2)\frac{\Delta R_{ij}^2}{R^2}, \tag{1.10}$$

and the distance to the beam is modified in a similar way, $d_{iB} = 1/p_{Ti}^2$.

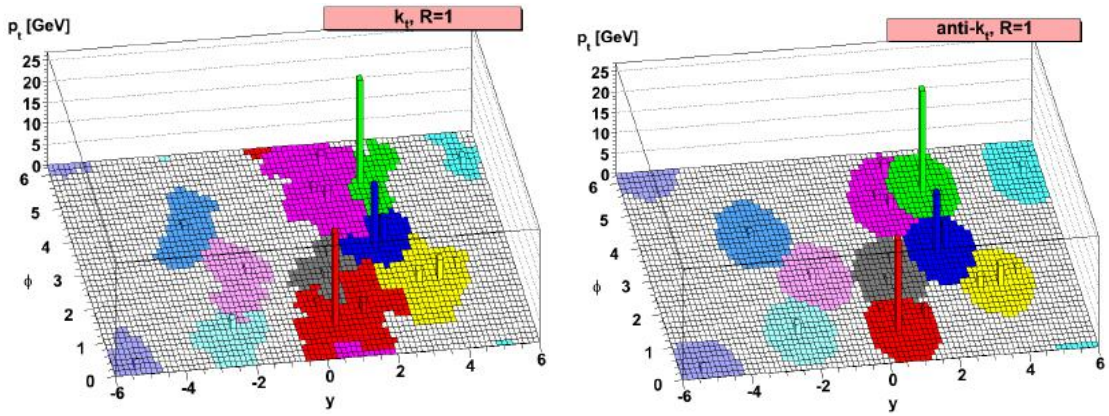
The functionality of the anti-k_{T} algorithm can be understood by considering an event with a few well separated hard particles (p_{T1}, p_{T2},...) and many soft particles. The d_{1i} between a hard and a soft particle will be determined by the transverse momentum of the hard one and the separation in the (η, ϕ) plane while the distance between two soft particles will instead be much larger. Therefore soft particles will tend to cluster with hard ones long before they cluster among themselves. If a hard particle has no hard neighbour within a distance 2R, then it will simply accumulate all the soft ones within a circle of radius R, resulting in a perfectly conical jet.

If another hard particle, with index 2, is present such that $R < \Delta_{12} < 2R$ then there will be two hard jets and it will be impossible for both to be perfectly conical; the particle with higher p_{T} will become perfectly conical and the other one will be partly conical, since it will miss the part overlapping with jet 1.

1.3.3 Comparison between algorithms

The behaviour of both algorithms is expected to be similar for hard particles, but it is expected to differ in the treatment of soft ones, particularly, in the regions around each of the final jets within which the soft particles are clustered. For the k_{T} algorithm that region depends somewhat on the specific set of soft particles which take part in the clustering, and the jagged borders of the jets are a consequence of the distribution of those particles. This behaviour means that the jet algorithm adapts its response to soft particles.

For the anti-k_{T} algorithm, the hard jets are all circular with a radius R, and only the softer jets have more complex shapes, which means that the jet algorithm is resilient to soft particles. An example of the different behaviours is shown in figure 1.4, where both algorithms are applied to a group of $\sim 10^4$ ghost particles (particles with infinitesimal p_{T}) in addition to those associated to an event from a pp collision.



(a) (b)

Figure 1.4: *Comparison between k_T and anti-k_T approaches to soft particles: the coloured areas around each jet are the regions in which soft particles are merged into a given jet [24].*

1.4 Top quark physics

The top quark is the heaviest of the quarks predicted by the Standard Model. It was proposed in 1973 by Makoto Kobayashi and Toshihide Maskawa as part of one of the models able to explain the observed CP violation in kaon decay [25] and experimentally discovered at Tevatron in the year 1995 [26]. The top quark is coloured, has electric charge +2/3 and forms an $SU(2)_L$ doublet with the bottom quark; because of these properties, it couples through the three interactions contained in the SM, strong, weak and electromagnetic.

The value of the top mass, as it is the case for all quark masses, is not predicted by the SM. It has been measured in various experiments since its discovery, which combined, yields a value of $m_{top} = 173.34 \pm 0.27$ (stat) ± 0.71 (syst) GeV [27]. The value of the top mass is very close to the electroweak symmetry breaking scale, and, therefore, it has a naturally strong coupling to the Higgs boson in the SM. It is because of this large coupling, for instance, that the Higgs mass can be predicted via precision measurements, and that the Higgs can be copiously produced at hadron colliders via top-loop mediated interactions to gluons. Its large mass can also be exploited in many scenarios that go beyond the standard model (BSM).

In addition to being massive, it also has a very short lifetime. In contrast to the lighter quarks, which are permanently confined in bound states with other quarks and antiquarks, the top quark decays so quickly that it does not have time to form such bound states. Thus, the top quark is free of many of the complications associated with the strong interactions.

1.4.1 Top quark pair production

Using the elements discussed in section 1.2 the production cross section for a top-antitop pair ($t\bar{t}$) is given by the expression:

$$\sigma_{pp\to t\bar{t}}(s, m_{top}) = \sum_{i,j} \int_0^1 dx_1 f_i(x_1, \mu_F^2) \int_0^1 dx_2 f_j(x_2, \mu_F^2) \sigma_{ij\to t\bar{t}}(x_1, x_2, s, m_{top}, \mu_R^2, \mu_F^2) \quad (1.11)$$

where $\sigma_{ij\to t\bar{t}}(x_1, x_2, s, m_{top}, \mu_R^2, \mu_F^2)$ is the partonic cross section and i,j runs (at lowest order) over the combinations of quarks and gluons of the types $q\bar{q}$ and gg. The lowest order Feynman diagrams that contribute to the partonic cross section are shown in figure 1.5.

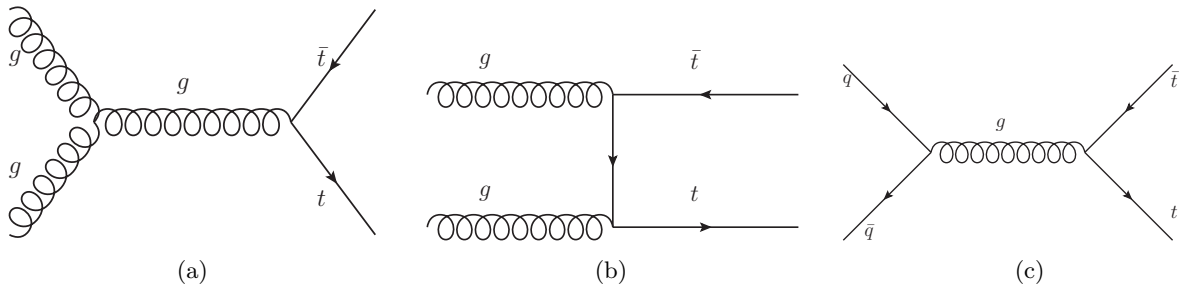


Figure 1.5: *Tree level Feynman diagrams for the top quark pair production at the LHC, both through gluon fusion (a and b) and quark-antiquark annihilation (c).*

The minimum amount of energy necessary to create a top quark pair is given by

$$\tilde{s} = 4m_{top}^2 \tag{1.12}$$

where $\tilde{s} = x_1 x_2 s$ is the effective centre-of-mass energy squared for the partonic process. At the LHC, small values of x_1 and x_2 are sufficient for $t\bar{t}$ production. Since the gluon PDF increases more steeply towards low x than the valence or even the sea-quark distributions (see figure 1.2) the production cross-section is dominated by gluon-gluon fusion, represented by figures 1.5a and 1.5b. In addition, the $t\bar{t}$ pairs are typically produced above the mass threshold due to the large available centre-of-mass energy, some of them with high transverse momentum (p_{T}).

The reference $t\bar{t}$ production cros section at a centre of mass energy of $\sqrt{s} = 7$ TeV used in this dissertation is $\sigma_{pp\to t\bar{t}} = 177^{+10}_{-11}$ pb for a top quark of mass 172.5 GeV. It has been calculated at next-to-next-to leading order (NNLO) in QCD including resummation of next-to-next-to-leading logarithmic (NNLL) soft gluon terms with top++2.0 [28–31]. The results were obtained with the MSTW2008 NNLO PDF set [32].

1.4.2 Top quark decay

In the SM, the masses and mixings of quarks arise from the Yukawa interactions with the Higgs condensate. In particular, the Cabbibo-Kobayashi-Maskawa (CKM) matrix [25, 33], a 3×3 unitary matrix, parameterises the strength and characteristics of all the flavour-changing processes in the SM. Except for its unitarity and the number of flavours, the CKM matrix is not fixed by the theory and its value has to be determined experimentally. The results of a global fit using all available measurements for the magnitudes of all nine CKM matrix elements are [21]:

$$\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.97427 \pm 0.00014 & 0.22536 \pm 0.00061 & 0.00355 \pm 0.00015 \\ 0.22522 \pm 0.00061 & 0.97343 \pm 0.00015 & 0.0414 \pm 0.0012 \\ 0.00886^{+0.00033}_{-0.00032} & 0.0405^{+0.0011}_{-0.0012} & 0.99914 \pm 0.00005 \end{pmatrix} \tag{1.13}$$

The decay modes of the top quark are governed by the last row of the CKM matrix elements and the large value of $|V_{tb}|$ with respect to the other elements. Therefore, the top decays almost exclusively into a W boson and a b quark ($t \to Wb$). The other possible decay modes ($t \to Ws$ and $t \to Wd$) are heavily suppresed by the magnitude of the corresponding CKM matrix elements. Henceforth, only the decay channel $t \to Wb$ will be considered.

The decay of the top quark is extremely fast, as mentioned above; the top lifetime is measured to be $\approx 0.5 \times 10^{-24}$ s [21], which is about 20 times faster than the typical strong interaction scale.

The W boson, with a mass of 80.385 ± 0.015 GeV, obtained by making a fit to all available measurements [21], also decays very rapidly. A W boson decays in about 1/3 of the time into

a charged lepton and a neutrino and all the three lepton flavours (electron, muon and tau) are produced at approximately the same rate. In the remaining 2/3 of the cases, the W boson decays into a quark-antiquark pair (of different flavours) and the abundance of a given pair is determined by the magnitude of the relevant CKM matrix element. The top quark is not accessible since its mass is higher than that of the W boson and the production of b quarks is heavily suppressed due to the values of $|V_{cb}|$ and $|V_{ub}|$.

Using this knowledge, three types of $t\bar{t}$ events can be distinguished from the parton configuration, that is, the products of the top and W boson decay processes, which are shown in figure 1.6:

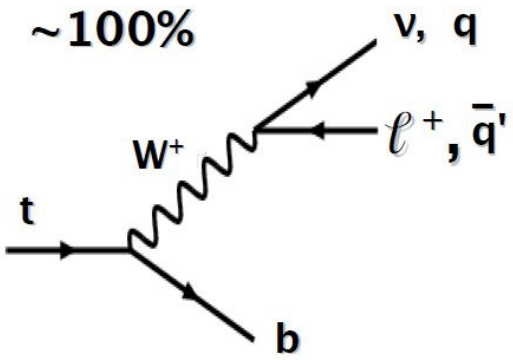


Figure 1.6: *Main decay channels of the top quark.*

- **fully-leptonic channel:** it represents about 1/9 of the $t\bar{t}$ events. Both W bosons decay into a charged-lepton-neutrino pair, resulting into an event with two charged leptons, two neutrinos and two b-quarks;
- **hadronic channel:** it represents about 4/9 of the $t\bar{t}$ decays. Both W bosons decay in quark-antiquark pairs, giving rise to an event with four light quarks and two b-quarks;
- **semi-leptonic channel:** it represents about 4/9 of the $t\bar{t}$ decays. One W-boson decays into a charged-lepton-neutrino pair, and the other one decays into a quark-antiquark pair, giving rise to an event with one charged lepton, one neutrino, two light quarks and two b-quarks.

A diagram showing the distribution of the different channels of $t\bar{t}$ decay is shown in figure 1.7. In this dissertation, only the semi-leptonic channel is considered. Furthermore, only final-states in which the charged lepton is an electron or positron are studied. The semi-leptonic channel is the preferred channel for the analysis as it provides a clear signature while at the same time keeping the possibility to measure all the decay products of the top quark that decayed hadronically and, therefore, to reconstruct the top quark directly.

1.5 Photon plus jets physics

1.5.1 Production of prompt photons in the LHC

The production of prompt photons in hadronic collisions proceeds through two mechanisms. In the first one, which is usually referred to as "direct", the photon takes part in the hard process and is well separated from any hadronic activity. In the other one, which is usually referred to

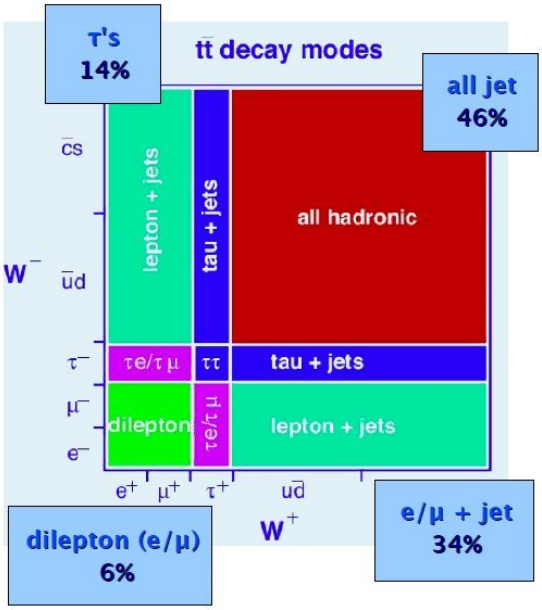


Figure 1.7: *Distribution of the different modes of $t\bar{t}$ decay.*

as "fragmentation", the photon results from the collinear fragmentation of a coloured high p_{T} parton and is accompanied by hadrons [34].

The LO contribution to direct production is given by the tree-level processes $q\bar{q} \to \gamma g$ and $gq\,(or\,\bar{q}) \to \gamma q\,(or\,\bar{q})$ and their corresponding Feynman diagrams are shown in figure 1.8. They are $O(\alpha\alpha_s)$.

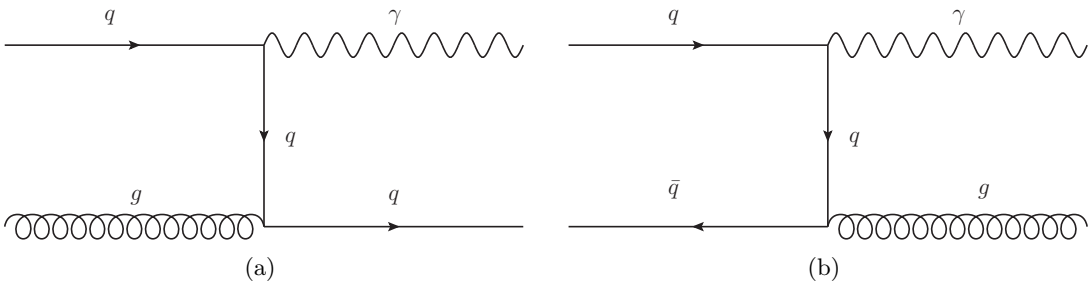


Figure 1.8: *LO Feynman diagrams for the direct contribution to prompt-photon production.*

From a technical point of view, the fragmentation contribution emerges from the calculation of the higher order corrections in the perturbative expansion in α_s, which at NLO includes the subprocesses $q\bar{q} \to \gamma gg$ and $gq\,(or\,\bar{q}) \to \gamma gq\,(or\,\bar{q})$ and virtual corrections to the direct production tree-level processes. Final state collinear singularities appear in any subprocess where a high p_{T} outgoing parton undergoes a cascade of successive collinear splittings together with the collinear emission of a photon. The higher order corrections to the cross-section can be split into a finite contribution, free of any singularity, to be added to the tree term so as to build the direct contribution and a fragmentation contribution involving these singularities.

The singularities can be factorised to all orders in α_s using the factorisation theorem described in section 1.2 and absorbed into fragmentation functions of a parton of type k to a photon $D_k^{\gamma}(z, \mu_f^2)$ where μ_f is some arbitrary fragmentation scale. This fragmentation functions have to be modeled in some way and/or constrained using experimental data. When the fragmentation scale is large with respect to $O(1)$ GeV, the typical hadronic scale, these functions behave roughly as $\alpha/\alpha_s(\mu_f^2)$ so that these contributions are of the same order as the leading

order contributions of the direct mechanism ($O(\alpha\alpha_s)$). The Feynman diagrams corresponding to the LO fragmentation component are shown in figure 1.9

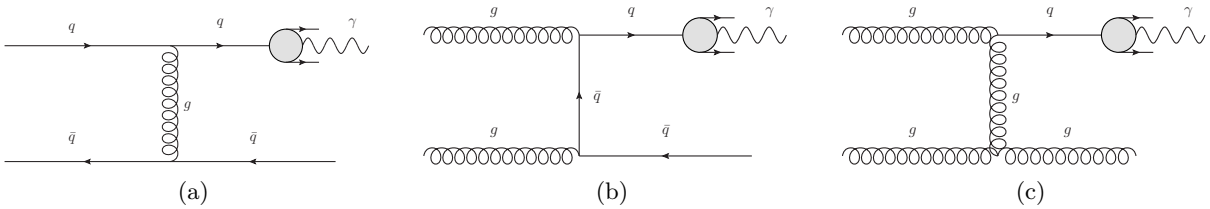


Figure 1.9: *LO Feynman diagrams for the fragmentation contribution to prompt-photon production.*

It is important to mention that the separation between direct and fragmentation contribution has no physical meaning beyond leading order. At leading order, the calculations for both direct and fragmentation processes are finite, and thus can be considered completely independent processes. At higher orders, it is necessary to consider both contributions at the same time in order to cancel the singularities. The splitting of the cross section between them is not unique and the fragmentation functions depend on the arbitrary factorisation scheme specifying which non-singular parts are factorised together with the collinear singularities in the fragmentation component and on the arbitrary fragmentation scale μ_f which is an unphysical parameter.

In general the inclusive photon-production cross-section can be written as

$$\sigma(p_\gamma) = \sigma^d(p_\gamma, \mu_R^2, \mu_F^2, \mu_f^2) + \sum_k \int_0^1 \frac{dz}{z} \sigma^f(p_\gamma/z, \mu_R^2, \mu_F^2, \mu_f^2) D_k^\gamma(z, \mu_f^2) \tag{1.14}$$

where σ^f and σ^d are the corresponding fragmentation and direct cross-sections. σ^f describes the production of a parton k (k=q,$\bar{q}$,g) in the hard collision and D_k^γ is the fragmentation function of a parton k into a photon. σ^d does not contain any fragmentation function and corresponds to the point-like coupling of the large p_T photon to a quark produced in the hard subprocess. Note that they are not true partonic cross-sections since they include the convolution with the parton distributions as discussed in section 1.2.

1.5.2 Isolation in prompt-photon production

In hadron colliders experiments, such as those at the LHC, it is not possible to perform inclusive photon measurements. The background from secondary photons coming from the decays of π^0,η, etc., overwhelms the signal by several orders of magnitude. To reject this background, the experimental selection of prompt photons requires isolation cuts. The isolation criterion usually employed by collider experiments, and used in this dissertation, is described in the following. A photon is said to be isolated if, in a cone of radius R in pseudo-rapidity and azimuthal angle around the photon direction, the amount of deposited hadronic transverse energy $E_\mathrm{T}^\mathrm{iso}$ is smaller than some value $E_\mathrm{T}^\mathrm{max}$ chosen for the specific analysis:

$$E_\mathrm{T}^\mathrm{iso} \leq E_T^{max} \text{ inside } (\eta - \eta_\gamma)^2 + (\phi - \phi_\gamma)^2 \leq R^2. \tag{1.15}$$

In addition to the rejection of the background of secondary photons, the isolation cuts also affect the prompt-photon cross section itself, in particular by reducing the fragmentation contribution.

Such an isolation criterion enforces additional phase-space restrictions; this implies that the cross section is no longer fully inclusive and, hence, that the factorised expression 1.14 is not

necessarily valid. In order to preserve factorisation, the isolation requirement has to fulfil the following requirements:

- **infrared safety**, which means that the cross section is insensitive to the momenta of arbitrarily soft particles;
- **collinear safety**, which means that when some final-state particles are produced collinearly, the cross section depends on their total momentum rather than on the momentum of each of them;
- **final state collinear factorisability**, which guarantees that all long-distance phenomena related to the low momentum fragmentation of the photon can be absorbed and factorised in the universal fragmentation functions;
- **initial state collinear factorisability**, which guarantees that the criteria does not spoil the absorption of initial-state collinear singularities into the hadron PDF.

It has been proven that the isolation criteria defined in 1.15 is both infrared and collinear safe and fulfils both factorisation properties to all orders in perturbation theory [35].

1.5.3 Photon-jet angular distribution

An observable which is expected to receive a distinctive contribution from the fragmentation component is the photon-jet angular distribution. At LO, corresponding to 2→2 kinematics, $\cos\theta^* \equiv \tanh(\Delta y/2)$ and Δy is the difference between the rapidities of two final-state particles. The variable θ^* coincides with the scattering angle in the centre-of-mass frame, and its distribution is sensitive to the spin of the exchanged particle. For processes dominated by t-channel gluon exchange, such as dijet production, the differential cross-section behaves as $(1-|\cos\theta^*|)^{-2}$ when $|\cos\theta^*| \to 1$. In contrast, processes dominated by t-channel quark exchange, such as $W/Z + jet$ production, are expected to have an asymptotic $(1-|\cos\theta^*|)^{-1}$ behaviour.

In photon plus jet production the direct-photon contribution, as seen in figure 1.8 is expected to exhibit a $(1-|\cos\theta^*|)^{-1}$ dependence when $|\cos\theta^*| \to 1$ while the fragmentation contribution, as seen in figure 1.9, is expected to exhibit the same dependence as dijet production, namely $(1-|\cos\theta^*|)^{-2}$. For both processes there are also s-channel distributions which are, however, non-singular when $|\cos\theta^*| \to 1$. As a result, a measurement of the cross-section for prompt-photon plus jet production as a function of $|\cos\theta^*|$ provides a handle on the relative contributions of the direct-photon and fragmentation components as well as the possibility to test the dominance of $t-$channel quark exchange.

1.5.4 Colour coherence effects in photon plus jets events

Colour coherence phenomena, in which the colour connection between partons influences the final-state topology, provide several important tests of QCD and, in particular, on how gluon radiation is included into QCD calculations beyond the leading order.

The most striking consequence of colour coherence phenomena in QCD is given by the suppression of soft radiation emission. This effect is understood as the result of destructive interference between amplitudes with soft gluons emitted by colour connected partons. The colour flow lines create "directional antennas" which behave approximately like standard dipoles, concentrating soft radiation mainly in the regions towards which the antennas are pointing; a simple example of colour flow between three partons is shown in figure 1.10.

Colour coherence effects have been traditionally studied in multijets events, both in Tevatron [36, 37] and LEP [38] experiments, where colour flow can be present both between initial

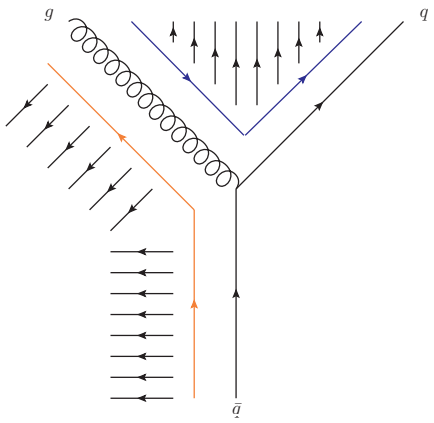


Figure 1.10: *Colour flow example between three partons in the final-state. The coloured arrows represent the colour flow and the small black arrows represent the regions where the soft radiation concentrates due to the antenna effect.*

and final-state particles (in the Tevatron case) or only between final-state particles (in the LEP case). Photon plus jets events provide an ideal environment to study this type of effects due to the fact that the final-state contains a coloured parton, which will partake on the colour flow and the creation of the aforementioned antennas and a colourless photon, which will not. It is expected that the production of soft radiation around the final-state parton would be enhanced in certain directions with respect to the production around the photon, in particular, in the direction defined by the antenna created between the final-state parton and one of the initial-state partons [39]. An example of a colour flow configuration in a photon plus jet event is shown in figure 1.11.

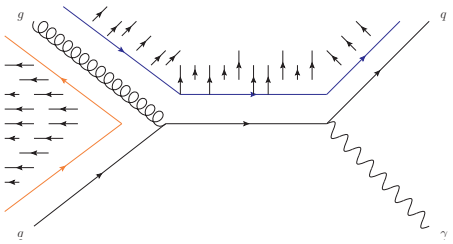


Figure 1.11: *Colour flow example in a photon plus jet event, the coloured arrows represent the colour flow and the small black arrows represent the regions where the soft radiation concentrates due to the antenna effect.*

Chapter 2

The LHC and the ATLAS detector

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC), built at CERN (Geneva) between 1998 and 2008, is the largest particle accelerator and collider in the world. The accelerator consists of a 27-kilometre ring of superconducting magnets with a number of accelerating structures to boost the energy of the particles along the way. The LHC was installed into the existing tunnel that was constructed between 1984 and 1989 for the CERN Large Electron Positron collider (LEP) [40, 41]. The LHC, a particle-particle collider, consists of two rings with counter-rotating beams, unlike particle-antiparticle colliders that can have both beams sharing the same ring. The LHC is therefore designed with separate magnet fields and vacuum chambers in the main arcs and with common sections only at the intersection regions (IR), where the experimental detectors are located.

Beams are kept on their circular paths by using 1232 dipole magnets, while an additional 392 quadrupole magnets are used to keep the beams focused. The electro-magnets are made of NbTi superconducting cables, cooled below 2 K by superfluid helium and operating at fields of 8.4 T. The external view of one of the dipoles is shown in figure 2.1.

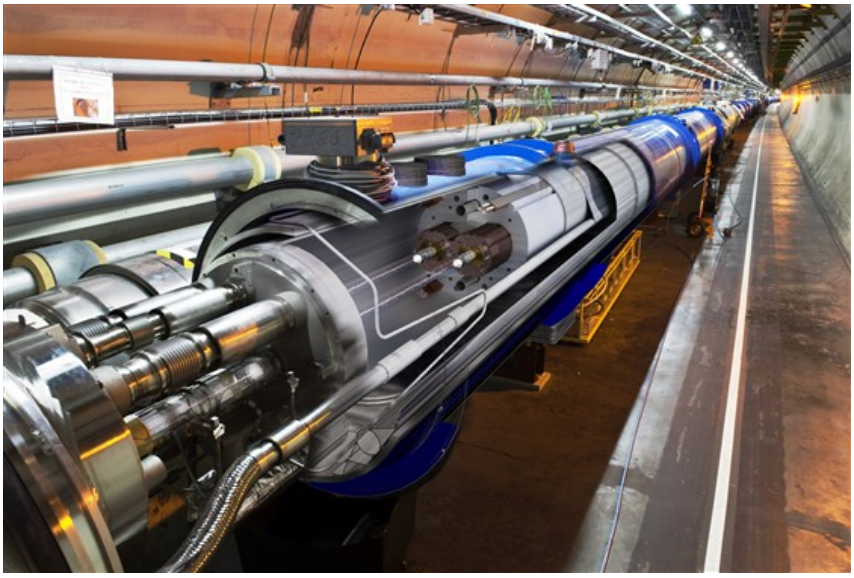


Figure 2.1: *Superconducting dipole magnet at the LHC.*

Prior to being injected into the main ring, particles are accelerated in several steps by a complex system of accelerators, shown in figure 2.2. First the Linear Particle Accelerator

(LINAC2) generates 50 MeV protons and feeds them into the Proton Synchrotron Booster (PSB) which then accelerates them to an energy of 1.4 GeV. Next, the beams are injected into the Proton Synchrotron (PS) where they reach an energy of 26 GeV. The beams are then accelerated using the Super Proton Synchrotron (SPS) which takes the protons to an energy of 450 GeV before injecting them into the LHC. Finally, the LHC uses a system of radio frequency cavities with a frequency of 400 MHz which is able to accelerate the 450 GeV beams to the 7 TeV design energy.

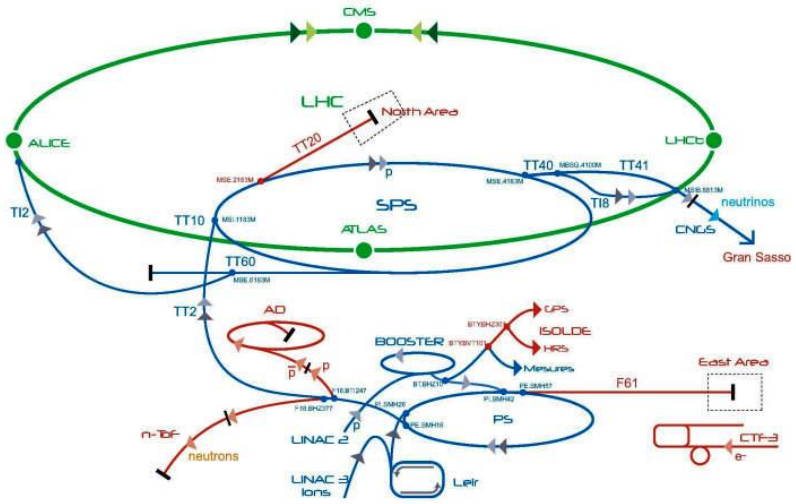


Figure 2.2: *Schematic view of the LHC accelerator ring and the system of accelerators used to deliver beams to the LHC.*

While operating in physics mode, with the intention of delivering collisions to the main experiments, the LHC accelerates and maintains beams of protons separated by 25 or 50 ns that intersect at four interaction points where the main detectors are located, ATLAS, CMS, ALICE and LHCb. ATLAS and CMS are general purpose detectors, designed for new physics searches and, specifically, for the search of the Higgs boson. The LHCb is a B-physics oriented detector and ALICE is a heavy ion experiment. The location of the experiments along the accelerator ring is shown in figure 2.2.

The number of events per second generated in the LHC collisions is given by

$$N_{event} = L\sigma_{event}, \tag{2.1}$$

where σ_{event} is the cross section for the event under study and L the machine luminosity. The machine luminosity depends only on the beam parameters and can be written for a Gaussian beam distribution as:

$$L = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi\epsilon_n \beta^*} F, \tag{2.2}$$

where N_b is the numbers of particles per bunch, n_b the number of bunches per beam, f_{rev} is the revolution frequency, γ_r the relativistic gamma factor, ϵ_n the normalised transverse beam emittance, β^* the beta function at the collision point, and F the geometric luminosity reduction factor due to the crossing angle at the interaction point:

$$F = \left(1 + \left(\frac{\theta_c \sigma_z}{2\sigma^*}\right)^2\right)^{-1/2}, \tag{2.3}$$

where θ_c is the full crossing angle at the interaction point (IP), σ_z the root mean square (RMS) of the bunch length and σ^* the transverse RMS of the beam size at the IP. The above expression assumes round beams, with $\sigma_z << \beta^*$, and with equal beam parameters for both beams. Since the cross sections of interesting physics process can be quite small, the LHC collisions requires both high beam energies and high beam intensities to be able to produce a high enough number of interesting events from said processes.

2.2 Brief timeline of the LHC and ATLAS data taking.

After finishing construction in 2008 the accelerator was expected to be operating at a centre of mass energy ($\sqrt{s}$) of 10 TeV by the end of 2009. Unfortunately, on September 19th, 2008 a magnet quench occurred in about 100 bending magnets, causing a loss of approximately 6 tonnes of liquid helium and breaking the vacuum conditions of the beam pipe. This event shaped the evolution of the LHC data taking during the following years. To run in full safety, the LHC operated up to $\sqrt{s} = 8$ TeV till the first long shutdown in 2013. A timeline of the LHC and ATLAS data-taking is summarised as follows:

- during 2009 the LHC first beams were circulated at $\sqrt{s} = 900$ GeV which allowed ATLAS to obtain approximately 9 μb^{-1} to use for calibration purposes;
- during 2010 the energy of the beams was gradually increased up to $\sqrt{s} = 7$ TeV, setting a new energy record and making the official start of the LHC physics program. The number of bunches and the bunch intensity of the machine were also increased up to 368 bunches and 1.2×10^{11} protons per bunch for a total maximum peak luminosity of $2.1 \times 10^{32}\ cm^{-2}s^{-1}$. A total integrated luminosity of 45 pb^{-1} was recorded by ATLAS. In figure 2.3a the integrated luminosity recorded by ATLAS in 2010 is shown;
- during 2011 the centre of mass energy was maintained at 7 TeV while the number of bunches was increased to 1380 and the bunch intensity to 1.45×10^{11} protons per bunch for a maximum peak luminosity of $3.7 \times 10^{33}\ cm^{-2}s^{-1}$. A total integrated luminosity of 5.08 fb^{-1} was recorded by ATLAS. In figure 2.3b the integrated luminosity recorded by ATLAS in 2011 is shown;
- during 2012 the centre of mass energy was increased to 8 TeV maintaining the number of bunches in 1380 and increasing the bunch intensity to 1.7×10^{11} protons per bunch for a maximum peak luminosity of $7.7 \times 10^{33}\ cm^{-2}s^{-1}$. A total integrated luminosity of 21.3 fb^{-1} was recorded by ATLAS. In figure 2.3c the integrated luminosity recorded by ATLAS in 2012 is shown;
- a long shutdown during 2013 and 2014 allowed repair work to be made in the machine to be able to reach the nominal luminosity of $1 \times 10^{34}\ cm^{-2}s^{-1}$ at a $\sqrt{s}$ of 13-14 TeV. It was possible to improve the superconducting circuit which was the cause of the 2008 incident. In addition, minor improvements and maintenance tasks to the experiments were performed;
- during the summer of 2015 the first collisions at $\sqrt{s} = 13$ TeV were delivered by the LHC and recorded by ATLAS. In figure 2.3d the integrated luminosity recorded by ATLAS in 2015 is shown.

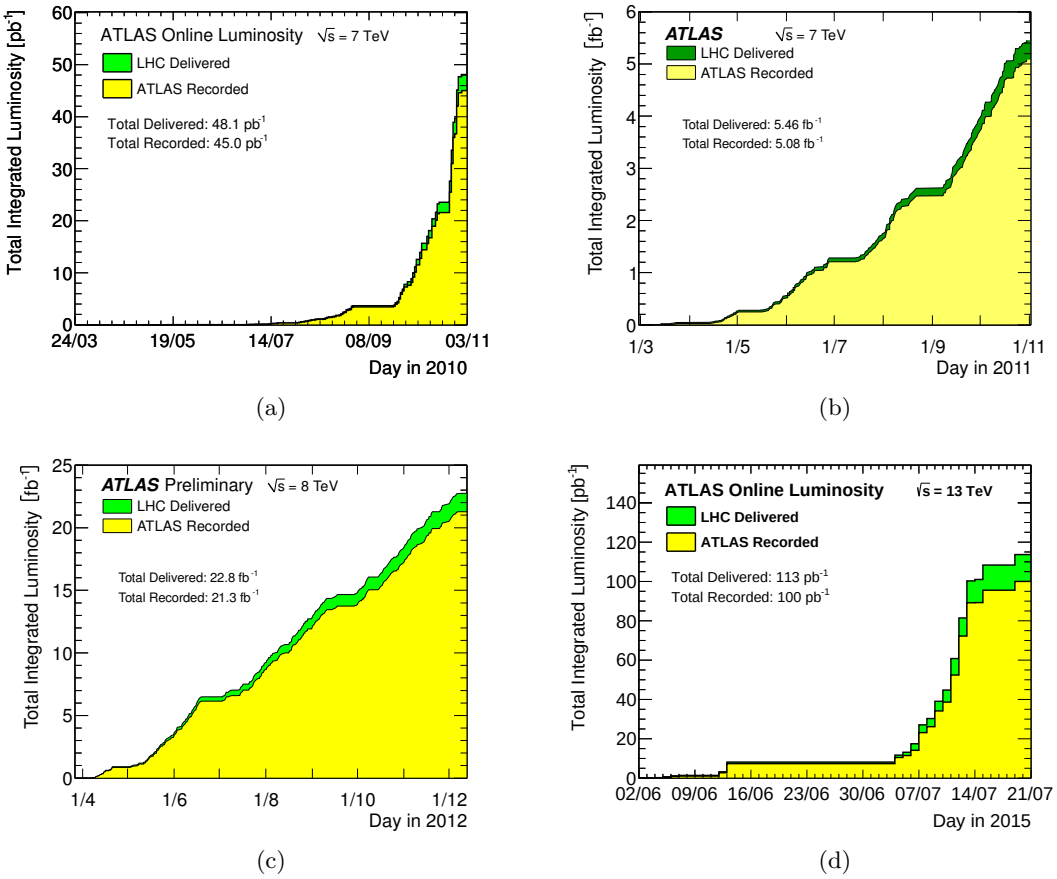


Figure 2.3: *Integrated luminosity delivered by the LHC and recorded by ATLAS during the different periods of Data taking.*

2.3 The ATLAS detector

ATLAS (A Toroidal Lhc ApparatuS) is a general purpose detector designed to fully use the discovery potential of the LHC. The detector, which is shown in figure 2.4 is a cylinder with a longitude of 44 m and a radius of 11 m and weighs approximately 7000 tons [42, 43]. Its different components are described in detail in the following sections.

The LHC large luminosity and resulting interaction rates are needed to study new physics and make precision measurements of QCD and electroweak processes, but they present a serious experimental difficulty as it implies that every candidate event for new physics will, on the average, be accompanied by 23 inelastic events per bunch crossing. The nature of proton-proton collisions imposes another difficulty. QCD jet production cross sections dominate over many rare processes, requiring the identification of various experimental signatures such as missing transverse energy (E_T^{miss}) or secondary vertices.

Due to these conditions and goals, a set of general requirements for the LHC detectors, ATLAS among them, is summarised as follows:

- the detectors require fast, radiation hard electronics and sensor elements. In addition, high detector granularity is needed to handle the particle fluxes and to reduce the influence of overlapping events;

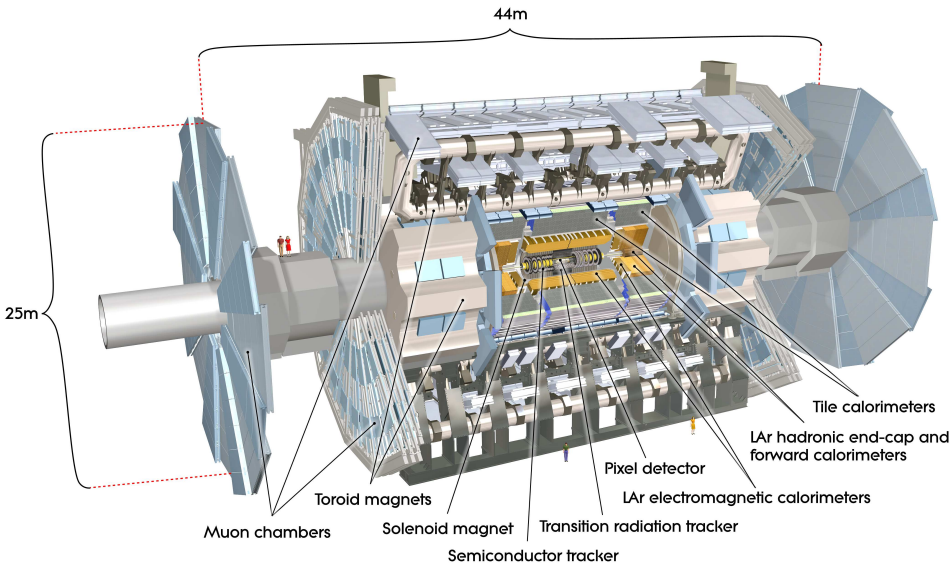


Figure 2.4: *Cut-away view of the ATLAS detector.*

- large acceptance in pseudorapidity (η) [1] with almost full azimuthal angle coverage is required;
- good charged-particle momentum resolution and reconstruction efficiency in the inner tracker are essential. For offline tagging of τ leptons and b-jets, vertex detectors close to the interaction region are required to observe secondary vertices;
- very good electromagnetic calorimetry for electron and photon identification and measurements, complemented by full-coverage hadronic calorimetry for accurate jet and missing transverse energy measurements are also important requirements;
- good muon identification and momentum resolution over a wide range of momenta and the ability to determine unambiguously the charge of high P_T muons are fundamental requirements;
- highly efficient triggering on low P_T objects with sufficient background rejection is a prerequisite to achieve an acceptable trigger rate for most physics processes of interest.

Following these guidelines the general performance goals for the different part of the ATLAS detector are shown in Table 2.1

2.3.1 Inner detector

The ATLAS Inner detector (ID) is a tracking detector, designed to measure the trajectory, also known as track, of charged particles. It consists of three independent sub-detectors and its layout is shown in figure 2.5. At inner radii, high resolution pattern recognition capabilities are available using discrete space-points from silicon pixel layers and stereo pairs of silicon microstrip (SCT) layers while at larger radii, the transition radiation tracker (TRT) comprises many layers

[1] ATLAS uses a coordinate system defined as follows: the nominal interaction point is defined as the origin of the coordinate system, while the beam direction defines the z-axis and the x-y plane is transverse to the beam direction. The azimuthal angle ϕ is measured around the beam axis, and the polar angle θ is the angle from the beam axis. The pseudorapidity is defined as $\eta = -log\tan\theta/2$ and the rapidity is defined as $y = 1/2log[E + p_z/(E - p_z)]$. The transverse momentum p_T, the transverse energy E_T and the missing transverse energy E_T^{miss} are defined in the x-y plane. The distance ΔR is defined in the $\eta - \phi$ plane.

Detector component	Required resolution	η coverage	
		Measurement	Trigger
Tracking	$\sigma_{P_t}/P_t = 0.05\%P_t \oplus 1\%$	± 2.5	
EM calorimetry	$\sigma_E/E = 10\%\sqrt{E} \oplus 0.7\%$	± 3.2	± 2.5
Hadronic calorimetry (jets)			
Barrel and end-cap	$\sigma_E/E = 50\%\sqrt{E} \oplus 3\%$	± 3.2	± 3.2
forward	$\sigma_E/E = 100\%\sqrt{E} \oplus 10\%$	$3.1 < \lvert\eta\rvert < 4.9$	$3.1 < \lvert\eta\rvert < 4.9$
Muon spectrometer	$\sigma_{P_t}/P_t = 10\%$ at $P_t = 1$ TeV	± 2.7	± 2.4

Table 2.1: *General performance goals of the ATLAS detector.*

of gaseous (Xenon-based gas mixture) straw tube elements interleaved with transition radiation material.

The ID is immersed in a solenoidal magnetic field of 2 T along the z-direction provided by the barrel solenoid, which allows the measurement of charged particles P_T using the curvature produced by the magnetic field. The solenoid is a single layer of Al stabilised Nb/Ti/Cu conductor cooled by liquid Helium at a temperature of $\sim 4.5K$ situated between the inner detector and the calorimeter.

To maintain an adequate noise performance even after radiation damage, the silicon detectors are kept at a low temperature (5 to 10^{o}C) while the TRT is designed to operate at room temperature.

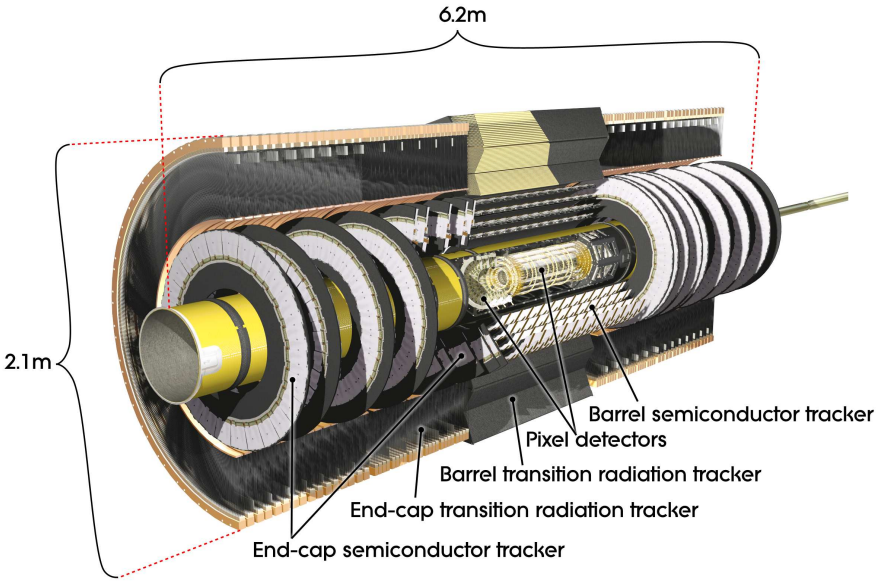


Figure 2.5: *Cut-away view of the ATLAS inner detector.*

- **Pixel detector:** it covers the region $|\eta| < 2.5$, segmented in R-ϕ and z. It has an intrinsic accuracy of 10μm (R-ϕ) and 115μm (z) in the barrel region and 10μm (R-ϕ) and 115μm (T) in the disks, placed in the end-cap regions perpendicular to the beam axis. The pixel detector has 80.4 million readout channels;

- **SCT:** it covers the same region in $|\eta|$ and comprises eight strip layers that are crossed by each track. In the barrel region it uses small-angle (40 mrad) stereo strips to measure both coordinates while in the end-cap regions, the detectors have a set of strips running radially and a set of stereo strips at small-angle. The intrinsic accuracies per module are

17μm (R-ϕ) and 580μm (z) in the barrel and 17μm (R-ϕ) and 580μm (R) in the end-caps. The total number of readout channels in the SCT is 6.3 million;

- **TRT:** it provides a large number of hits (typically 36 per track), measured with 4mm diameter straw tubes, and enables track following up to $|\eta| = 2.0$. The TRT only provides R–ϕ information, and has an intrinsic accuracy of 130μm per straw. In the barrel region, the straws are parallel to the beam while in the end-cap region they are arranged radially in wheels. The total number of TRT readout channels is 351000.

The combination of precision trackers at small radii with the TRT at larger radius gives very robust pattern recognition and high precision in both R-ϕ and z coordinates. The straw hits at the outer radius contribute significantly to the momentum measurement, since the lower precision per point is compensated by the large number of measurements.

The inner detector provides tracking measurements in a range matched by the precision measurements of the electromagnetic calorimeter. The electron identification capabilities are enhanced by the detection of transition-radiation photons in the straw tubes. The semiconductor trackers also allow impact parameter measurements and vertexing for heavy-flavour and τ-lepton tagging. The secondary vertex measurement performance is enhanced by the innermost layer of pixels.

2.3.2 Calorimetry

The calorimeters of ATLAS are designed to provide energy measurements of the particles created in the collisions. They are sampling calorimeters, such that particles are absorbed by high density materials and their energy is measured by periodically sampling the shape of the resulting shower.

The calorimeters cover the range $|\eta| < 4.9$ using different techniques suited to the widely varying requirements of the physics processes of interest and of the radiation environment over this large η-range. In the η region matched to the inner detector, the fine granularity of the electromagnetic (EM) calorimeter is ideally suited for precision measurements of electrons and photons, the coarser granularity of the rest of the calorimeter is sufficient to satisfy the physics requirements for jet reconstruction and E_T^{miss} measurement.

They must provide good containment for electromagnetic and hadronic showers and must also limit punch-through into the muon system. The total thickness of the EM calorimeter is larger than 22 radiation lengths in the barrel and larger than 24 radiation lengths in the end-caps.

The calorimeter system is divided into 4 main subsystems, the Liquid Argon (LAr) electromagnetic calorimeter, the Tile calorimeter, the LAr hadronic end-cap calorimeter (HEC) and the LAr forward calorimeter (FCal), that are described in the following. Figure 2.6 shows the calorimetry layout in ATLAS.

- **LAr EM calorimeter:** it provides energy measurement for photons and electrons. It is divided into a barrel part ($|\eta| < 1.475$) and two end-cap components ($1.375 < |\eta| < 3.2$) each housed in their own cryostat kept at ~ 88 K. The position of the central solenoid in front of the EM calorimeter demands optimisation of the material to achieve the desired performance. As a consequence, the central solenoid and the LAr EM calorimeter share a common vacuum vessel. The barrel calorimeter consists of two identical half-barrels, separated by a small gap. Each end-cap calorimeter is mechanically divided into two coaxial wheels, an outer wheel covering the region $1.375 < |\eta| < 2.5$ and an inner wheel covering the region $2.5 < |\eta| < 3.2$.

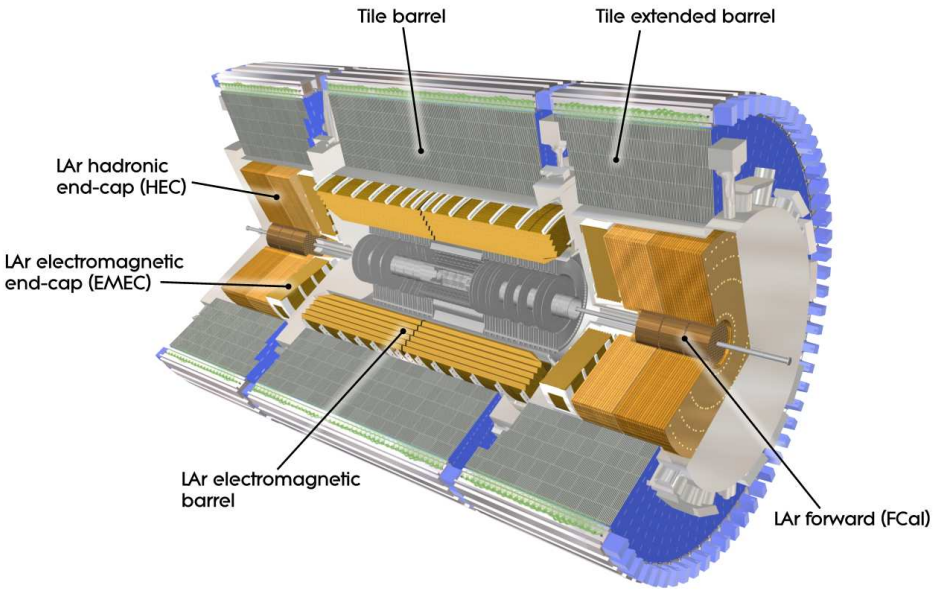


Figure 2.6: *Cut-away view of the ATLAS calorimeter system.*

It uses liquid Argon as the active detector medium, which was chosen for its intrinsic linear behaviour, its stability of response over time and its intrinsic radiation-hardness. It was constructed with accordion shape absorbers (made of lead) and electrodes (made of kapton). This geometry provides complete ϕ symmetry without azimuthal cracks. In the barrel the accordion waves are axial and run in ϕ, and the folding angles on the waves vary with radius to keep the liquid Argon gap constant while in the end-caps the waves are parallel to the radial direction and run axially. This geometry allows for a very uniform performance in terms of linearity and resolution as a function of ϕ. The calorimeter is divided into three active sections in depth for the region devoted to precision physics $|\eta| < 2.5$ and two for the rest of the coverage. In the region of $|\eta| < 1.8$ a pre-sampler detector is used to correct for the energy lost by electrons and photons upstream of the calorimeter;

- **Tile calorimeter:** it is placed directly outside the EM calorimeter envelope. Its barrel covers the region $|\eta| < 1.0$ and its two extended barrels the range $0.8 < |\eta| < 1.7$. It uses steel as the absorber and scintillating tiles as the active material. It is divided azimuthally into 64 modules and segmented in depth in three layers. The total detector thickness at the outer edge of the tile-instrumented region is 9.7 interaction lengths (λ) at $\eta = 0$;

- **HEC:** it consists of two independent wheels per end-cap, located directly behind the end-cap EM calorimeter and sharing the same LAr cryostats. The HEC extends covers the region $1.5 < |\eta| < 3.2$, overlapping with both the Tile and the FCal calorimeters. Each wheel is built from 32 identical wedge-shaped models, assembled with fixtures at the periphery and at the central bore. Each wheel is divided into two segments in depth, for a total of four layers per end-cap. The copper plates are interleaved with 8.5 mm LAr gaps, providing the active medium;

- **FCal:** it is integrated into the end-cap cryostats, as this provides clear benefits in terms of uniformity of the calorimetric coverage as well as reduced radiation background levels in the muon spectrometer. In order to reduce the amount of neutron albedo in the inner detector cavity, the front face of the FCal is recessed by about 1.2 m with respect to the EM calorimeter front face. This severely limits the depth of the calorimeter and therefore

calls for a high density design. The FCal is approximately 10 interaction lengths deep, and consist of three modules in each end-cap: The first, made of copper, is optimised for EM measurements, while the other two, made of tungsten, measure predominantly the energy of hadronic interactions. Each module consist of a metal matrix, with regularly spaced longitudinal channels filled with the electrode structure consisting of concentric rods and tubes parallel to the beam axis. The LAr in the gap between the rod and the tube is the sensitive medium. This geometry allows for excellent control of the gaps, which are as small as 0.25 mm in the first section in order to be able to withstand the particle fluxes in that region.

2.3.3 Muon system

The muon spectrometer in the ATLAS detector is a very large system and it is shown in figure 2.7. It is based on the magnetic deflection of muon tracks in the large superconducting air-core toroid magnets, instrumented with separate trigger and high-precision tracking chambers. Over the range $|\eta| < 1.4$, magnetic bending is provided by the large barrel toroid while for $1.6 < |\eta| < 2.7$ muons are bent by two smaller end-cap magnets inserted into both ends of the barrel toroid. In the transition region $1.4 < |\eta| < 1.6$ magnetic deflection is provided by a combination of both fields. This magnet configuration provides a field which is mostly orthogonal to the muon trajectories. As in the case of the solenoid, the toroid magnets are made of Al stabilised Nb/Ti/Cu conductor cooled by liquid Helium at a temperature of ~ 4.5 K and provide a magnetic field of approximately 1 T in the barrel region and 0.5 T in the end-cap region.

In the barrel region, tracks are measured in chambers arranged in three cylindrical layers around the beam axis while in the transition and end-cap regions the chambers are installed in planes perpendicular to the beam, also in three layers. There are two types of chambers that provide precision tracking:

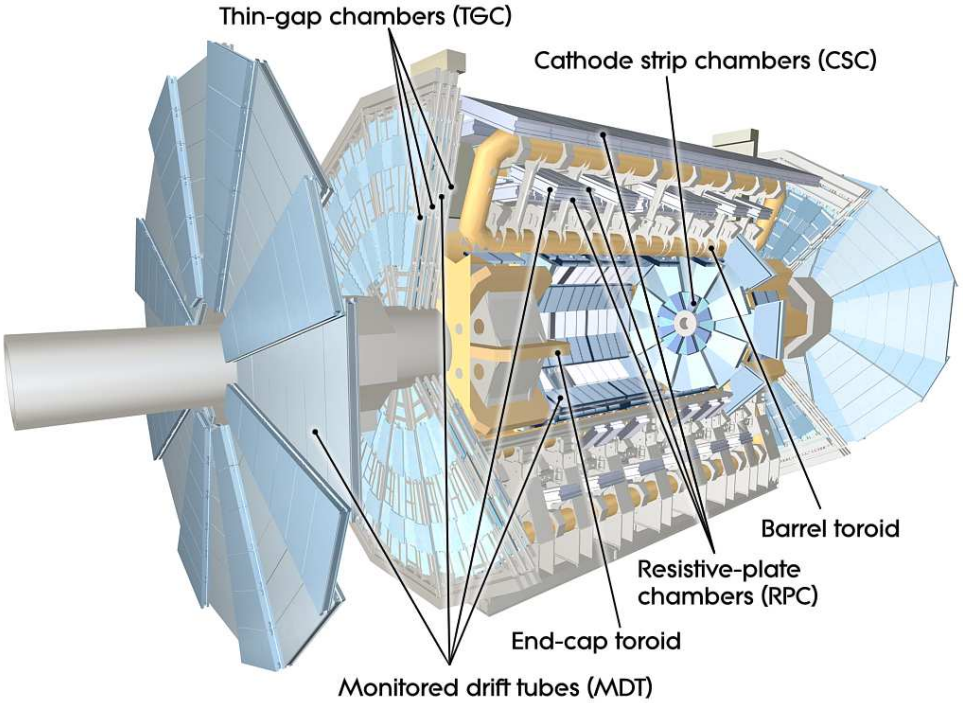


Figure 2.7: *Cut-away view of the ATLAS muon system.*

- **monitored Drift Tubes (MDT)** provide coverage in the region $|\eta| < 2.7$ ($|\eta| < 2.0$ for the innermost layer). It consists of three to eight layers of drift tubes, using Ar/CO_2 gas operated at an absolute pressure of 3 bar, arranged along the ϕ direction, that is, the centre points of the tubes are tangential to circles around the beam axis. It has an average resolution of 80 μm per tube or about 35 μm per chamber;

- **the Cathode-Strip chambers (CSC)** are used in the inner-layer in the forward region ($2.0 < |\eta| < 2.7$) due to their higher rate capability and time resolution. They are multi-wire proportional chambers with cathode planes segmented into strips in orthogonal directions. This allows both coordinates to be measured from the induced-charge distribution. The resolution of a chamber is 40 μm in the bending plane and about 5 mm in the transverse plane.

The precision-tracking chambers are complemented by a system of fast triggers chambers capable of delivering track information within a few tens of nanoseconds. **Resistive Plate Chambers (RPC)** are used in the barrel region ($|\eta| < 1.05$) and **Thin Gap Chambers (TGC)** are used in the end-cap ($1.05 < |\eta| < 2.4$) . They measure both coordinates of the track, one in the bending (η) plane and one in the non-bending (ϕ) plane which complements the bending plane high precision measurement of the MDT and CSC.

2.3.4 Forward detectors

Three smaller detector systems cover the ATLAS forward region. The main function of the first two systems is to determine the luminosity delivered to ATLAS. At $\pm$17 m from the interaction point lies **LUCID (LUminosity measurement using Cerenkov Integrating Detector)**. It detects inelastic scattering in the forward direction, and is the main online relative-luminosity monitor for ATLAS. The second detector is **ALFA (Absolute Luminosity For ATLAS)**. Located at $\pm$240 m, it consists of scintillating fibre trackers located inside roman pots which are designed to approach as close as 1 mm to the beam. The third system is the **Zero-Degree Calorimeter (ZDC)** which plays a key role in determining the centrality of heavy-ion collisions and is located at $\pm$ 140m from the interaction point. The ZDC modules consist of layers of alternating quartz rods and tungsten plates which measure neutral particles at $|\eta| \geq 8.2$.

2.3.5 Trigger system

The trigger system is required to reduce the initial bunch crossing rate of 40Mhz down to a manageable 100-200 Hz while retaining and recording the most interesting events, especially those with small cross section and low production rate. In figure 2.8 an estimation of the rates of several interesting physics processes is shown.

The trigger system consists of three levels of event selection:

- **Level-1 (L1) trigger:** it is implemented using custom-made electronics. It searches for signatures from high P_T muons, electron/photons, jets and τ leptons and it also selects events with large E_T^{miss} and large total transverse energy. It uses reduced granularity information from a subset of detectors. The maximum L1 accept rate which the detector readout system can handle is 75 KHz (upgraded to 100 KHz in 2015) and the L1 decision must reach the front-end electronics within μs after the bunch-crossing with which it is associated;

- **Level-2 (L2) trigger:** it is seeded by Regions-of-Interest (RoI's). These are regions of the detector where the L1 trigger has identified possible trigger objects within the event.

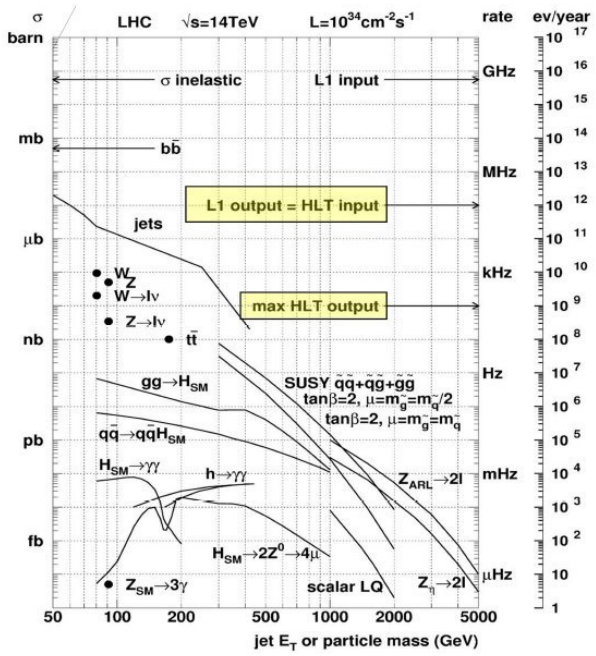


Figure 2.8: *Estimated rates of several physics processes and trigger levels.*

The L2 trigger uses RoI information on coordinates, energy and type of signatures to limit the amount of data which must be transferred from the detector readout. The L2 trigger reduces the event rate to below 3.5 KHz, whith an average event processing time of approximately 40 ms;

- **Event Filter (EF):** it uses offline analysis procedures on fully-built events to further select events down to a rate which can be recorded for subsequent offline analysis. It reduces the event rate to approximately 200 Hz, with an average event processing time on the order of 4 s.

The Level-1 Calorimeter trigger The L1 calorimeter trigger decision is based on dedicated analogue trigger signals provided by the ATLAS calorimeters independently from the signals read out and used at the high level trigger (HLT, composed of the L2 and the EF trigger) and offline reconstruction. Rather than using the full granularity of the calorimeter, the L1 decision is based on the information from analogue sums of calorimeter elements within projective regions, called trigger towers. The trigger towers have a size of approximately $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ in the central part of the calorimeter, $|\eta| < 2.5$ and are larger and less regular in the more forward region. Electromagnetic and hadronic calorimeters have separate trigger towers. The 7168 analogue inputs must first be digitised and then associated to a particular LHC bunch-crossing. Once digital transverse energies per LHC bunch crossing are formed, two separate processor systems, working in parallel, run the trigger algorithms. One system, **the cluster processor**, uses the full L1 trigger granularity information to look for small localised clusters typical of electrons, photons or tau particles. The other, **the jet and energy-sum processor**, uses 2×2 sums of trigger towers, called jet elements, to identify jet candidates and form global transverse energy sums, missing transverse energy, total transverse energy and jet-sum transverse energy. Details on the trigger process for both electrons and photons can be found in section 4.1.

Chapter 3

Monte Carlo simulations

To first approximation, high-energy processes have a simple structure at the level of the interactions between the fundamental objects of nature, i.e. quarks, leptons and gauge bosons. Corrections to this picture are of three main types:

- bremsstrahlung-type corrections, i.e. final-state particles can emit additional particles by branchings such as $e \to e\gamma$ or $q \to qg$. This emission is particularly prolific in the case of QCD emissions off quarks and gluons and a single initial parton may give rise to a whole bunch of partons in the final state. These corrections may be included to arbitrarily high orders using a probabilistic approach. Alternatively, exact fixed order calculations of bremsstrahlung corrections may be carried out order by order in perturbation theory, but the calculations then become prohibitively complicated very rapidly;
- it is also necessary to consider true higher order corrections, which involve a combination of loop graphs and the soft parts of the bremsstrahlung graphs mentioned above, a combination needed to cancel some divergences. The necessary perturbative calculations are usually very difficult;
- finally, quarks and gluons are confined. For quarks and gluons the perturbative language used to describe the short-distance interactions have to be complemented with the structure of the incoming hadrons and also the hadronisation process, wherein the coloured partons are transformed into jets of colourless hadrons.

The simple structure has now become considerably more complex, instead of maybe two final-state partons, the process has a large number of final-state particles. The original physics is not gone, but the skeleton process has been dressed up and is no longer directly visible. A direct comparison between theory and experiment is therefore very much complicated.

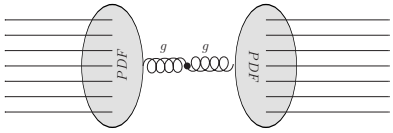
In Monte Carlo (MC) simulations, the objective is to generate events as detailed as could be observed by a perfect detector. The complex problem described above is factorised into a number of components, each of which can be handled with reasonable accuracy. The output of a MC generator should be in the form of "events" with the same average behaviour and the same fluctuations as in real data. In the data, fluctuations arise from the quantum mechanics of the underlying physics mechanism; in generators, MC techniques are used to select all relevant variables according to the desired probability distributions, and thereby ensure a (quasi-) randomness in the final state.

MC simulations are extremely useful for developing new analysis methods in the absence of real data and to understand the feasibility of proposed measurements. Under the assumption of the existence of a given signal, measurements depend intrinsically on theoretical models. Subtle differences in such models can affect measurements in various ways, but the effect is

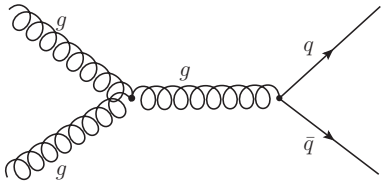
often unquantifiable without MC simulations; for this reason MC becomes an essential tool for discriminating theoretical models in terms of experimental observables as the simulations are entirely under the user's control. In addition, MC simulations are essential to be able to quantify detectors effects. By being able to access event information before and after it passes through the detector, it is possible to correct the data or to correct theoretical calculations to a level where it can be compared with the data.

A MC simulation of a proton-proton collision can be divided into the following general steps:

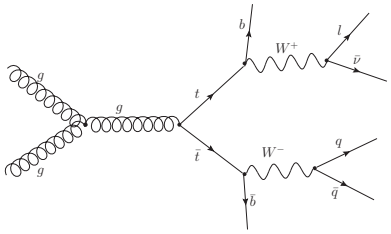
- as explained in section 1.2, PDF sets are used to correctly simulate the distribution of the incoming partons from the proton beams. The PDFs are modelled using experimental data and input to the MC program;



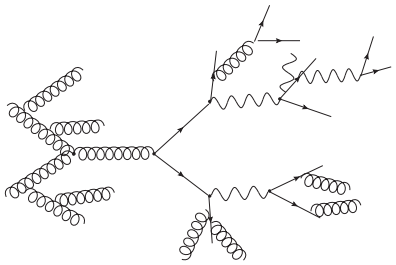
- one incoming parton from each proton enters the hard-scattering process. The hard-scattering process is then generated according to the matrix elements (ME). Usually only theoretical calculations at leading order (LO) are used, but this is not always the case;



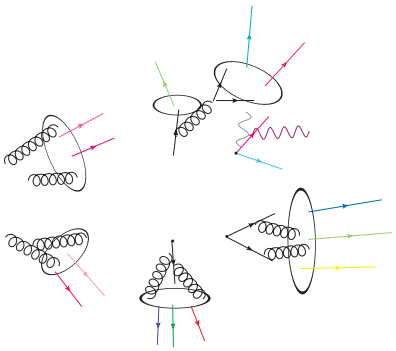
- short lived particles produced in the hard-scattering are decayed. This is achieved by using either experimental measurements of the width of the particles involved or calculated theoretically;



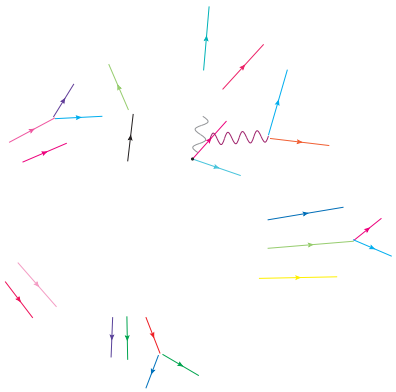
- initial and final state radiation is calculated, by branching the incoming and outgoing partons, and included in the process;



- the interactions of the spectator partons from the incoming protons is calculated and included; in addition, the remnant may have an internal structure and a net colour charge that is connected to the rest of the initial-state. The resulting particles form what is also known as the underlying event;
- additional interactions of pairs of protons in the same crossing are taken into account. These interactions are generally soft but can generate extra semi-hard events. This is also known as pileup;
- at this stage, the event is made up of partons and it is known as the parton-level of the MC simulation. From these final-state partons, the hadronisation process is started. This process is not well understood theoretically since it involves soft interactions that cannot be described in a perturbative approach to QCD and has to be modelled;

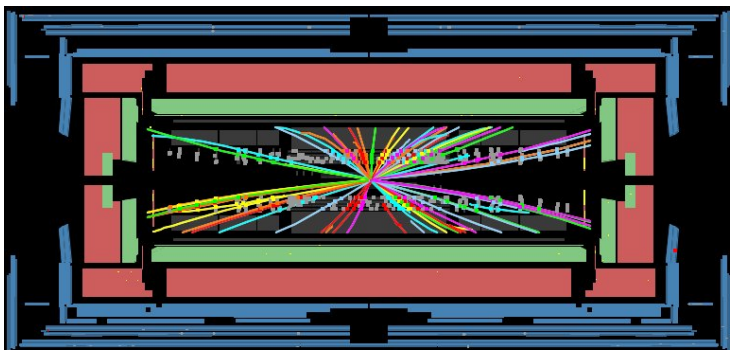


- the short-lived hadrons are decayed into more stable particles. The set of particles after this process is known as the hadron-level or particle-level of the MC simulation. The definition of stable particles is somewhat arbitrary but in the context of high-energy collisions a particle is considered stable if it has a long lifetime so as to interact with the detector material before decaying. In the ATLAS MC simulations a particle is considered stable if its lifetime is longer than 10 ps;



- this set of stable particles is provided as input to the detector simulation. This step is an extensive one in which:
  - the interaction of the particles with any magnetic field present is modelled;
  - the interaction of particles with matter is also simulated; the progress through the different layers of the detector is modelled with high precision;
  - the response of each subdetector component is also simulated. This is done to be able to provide an output signal as close as possible to that recorded in data. The resulting signals are processed by the same software as the data.

At the end of this step, a set of variables similar to those obtained with real collisions are made available to be used in the analysis.



3.1 Comparison between MC programs

The exact implementation of several of the previously detailed steps can be very different for each MC simulation program. In addition, multiple programs are often combined to generate samples of events, each taking care of one part of the simulation. In the following sections, additional details are given on how some of the most important steps are implemented by the different MC programs required to generate the samples used in this dissertation.

3.1.1 Hard scattering

Theoretical calculations of the hard scattering lie at the start of any MC generator. Once the interacting parton has been chosen from the PDF, the characteristics of the hard scattering have to be generated. This not only includes the calculation of the matrix element itself, but also providing the initial configuration of final-state partons that will be used event by event to generate the full simulation.

- in Pythia [44], only $2 \to 2$ or $2 \to 1$ processes are simulated. The matrix elements are only LO. It was used to generate prompt photon production samples and dijet samples;
- Sherpa [45] is able to calculate $2 \to n$ matrix elements at tree level, i.e. it is capable of including additional partons to the final state directly in the matrix element calculation. It was used to generate prompt photon production samples with up to 3 additional partons in the final state;
- Alpgen [46] is also able to calculate $2 \to n$ matrix elements at tree level, including additional partons to the final state. It was used to generate W+jets samples with up to 5 additional partons in the final state;
- Mc@nlo [47] is able to calculate NLO matrix elements, including real and virtual corrections. It was used to generate $t\bar{t}$ and single top quark production samples;
- AcerMC [48] is able to calculate LO matrix elements for high multiplicity final states. It was used to generate single top quark production samples.

3.1.2 Initial- and final-state radiation

From the initial and final-state particles obtained in the previous step, additional QCD and QED radiation is obtained by using what is called the parton shower approach. In this approach

the additional radiation is interpreted as a series of independent splitting processes $a \to bc$, such as $q \to qg$ or $g \to gg$. The evolution of the parton shower is governed by the DGLAP equations, deciding when and how each splitting takes place. Energy momentum conservation is explicit branching by branching.

The evolution of the parton shower is characterised by a virtuality scale Q^2, which gives an approximate sense of time ordering to the cascade. Final state showers are time-like, i.e. partons have $m^2 = E^2 - p^2 \geq 0$. Starting from some maximum scale Q^2_{max}, an original parton is evolved downwards in Q^2 until a branching occurs, distributing its energy among its daughters. The process of successive branchings is cut-off at some lower scale Q_0, which is typically around 1 GeV for QCD branchings.

Initial state showers are space-like, this means that in the sequence of branchings $a \to bc$ that lead up from the shower initiator to the hard interaction, particles a and b have $m^2 = E^2 - p^2 < 0$. It is handled within the backwards evolution scheme, that is, starting from the two incoming partons at the hard interaction, showers are traced "backwards in time" (in descending values of Q^2), towards the two shower initiators. In this approach, the choice of the hard scattering is based on the use of evolved parton distributions, which means that the inclusive effects of initial-state radiation are already included.

The implementation of the parton shower approach in the three MC programs used with that purpose in this dissertation differ mainly on the choice of the evolution scale Q^2:

- in PYTHIA the evolution scale is defined as $Q^2 = P_T^2 = z(1-z)m^2$ for final state radiation, where z describes the fraction of the original energy taken by one of the daughters ($E_b = zE_a$, $E_c = (1-z)E_a$), and as $Q^2 = P^2 = -(1-z)m^2$ for initial state radiation. This gives rise to P_T ordered branchings. In QCD showers, corrections to the leading-log picture, called coherence effects, lead to an ordering of subsequent emissions in terms of decreasing angles. The P_T ordered shower automatically leads to the correct angular ordering;

- in HERWIG [49], the evolution scale is defined by $Q^2 = 2E_a^2(1 - cos\theta_{bc})$ where θ_{bc} is the angle between the daughter partons. This choice of evolution scale leads automatically to angular ordering, in which the maximum subsequent angular opening is always the one from the previous branching;

- in SHERPA, the evolution scale is defined as $Q^2 = p^2 - m^2$. Color coherence during evolution is taken into account by an explicit angular veto, which means that a branching is rejected if the opening angle of the emission is larger than the one of the previous branching.

3.1.3 Matching between matrix elements and parton showers

When using ME and parton showers at the same time it is necessary to match them in a way that avoids double counting. This arises if a specific configuration has been generated already in the ME and is generated again using the parton shower approach. A matching scheme defines, on an event-by-event basis, which of the two paths should be followed according to the kinematics. Although they are implemented in different ways, all matching schemes of LO MC programs follow a similar strategy:

- a jet observable is defined and all relevant cross sections for the processes to be generated with additional n-jets in the final state are evaluated, ($pp \to X +$ njets with n $= 0, 1, 2...$);

- hard parton samples are produced with a probability proportional to the respective cross section, in a corresponding kinematic configuration following the matrix element;

- the individual configurations are accepted or rejected with a dynamical, kinematics-dependent probability. In case the event is rejected, the previous step is repeated, possibly with a new number of jets;
- the parton shower is invoked with suitable initial conditions. In all cases, the parton shower is constrained not to produce any extra jets, i.e. configurations that would fall into the realm of matrix elements with a higher jet multiplicity are vetoed in the parton shower step.

In NLO MC programs such as MC@NLO the NLO calculations include subtraction terms that correspond to the LO contributions of the parton shower. That is, the NLO calculation subtracts the LO parton shower contribution inside the calculation, avoiding any possible double counting.

3.1.4 Hadronisation

The hadronisation, in which colourless hadrons are constructed from the coloured final state particles exiting the parton shower, is based on the *Lund string model* [50] in the case of PYTHIA and on the *cluster model* [51] in the case of SHERPA and HERWIG:

- **Lund model:** it is a rather complicated model based on the idea that the long distance QCD interaction can be described by colour strings. These colour strings connect the partons and govern the transfer of momentum between them. This strings can, and will, be fragmented, and they gain energy as the partons get separated. This fragmentation is the source of new partons, that at the end of the process couple with the primary ones to create hadrons. The four-momentum of the hadrons is calculated using information both from the final partons and the connecting strings. A large number of additional parameters is used to fine-tune the relative production of particles such as strange, pseudoscalar, vector mesons, etc., by using data;
- **cluster model:** it is local in colour and independent of the hard process and the energy. This model is based of the preconfinement property of QCD. After the perturbative parton showering, all outgoing gluons are split non-perturbatively into light quark-antiquark or diquark-antidiquark pairs. At this point, the event consists of a set of outgoing quarks and antiquarks. In the limit of a large number of colours, each final state colour line can be followed from a quark/anti-diquark to an antiquark/diquark with which it can form a colour-singlet cluster. By virtue of pre-confinement, these clusters have a distribution of mass and spatial size that peaks at low values, falls rapidly for large cluster masses and sizes, and is asymptotically independent of the hard subprocess type and scale. The clusters thus formed are fragmented intro hadrons. If a cluster is too light to decay into two hadrons, it is taken to represent the lightest single hadron of its flavour and its mass is shifted to the appropriate value by an exchange of 4-momentum between the neighbouring clusters.

3.1.5 Underlying event

The underlying event is modelled by the introduction of Multiple Parton Interactions (MPI). Additional interactions are considered between the spectator partons of the hard process. The basic idea of such a model is to postulate the probability distribution of multiple scatterings. This is given as a function of the non-diffractive cross section σ_{ND}, the hard perturbative cross

section σ_{hard} and the outgoing transverse momentum p_{T} in the scattering by

$$p(p_{\text{T}}, b) = f_c f(b) \frac{1}{\sigma_{\text{ND}}} \frac{d\sigma_{\text{hard}}}{dp_{\text{T}}}, \tag{3.1}$$

where the prefactors f_c (normalisation) and $f(b)$ (proton shape function) incorporate an additional impact-parameter (b) dependence of the distribution on an event-by-event basis. In order to obtain a well-defined differential cross section in perturbation theory, a minimum p_{T} scale is introduced below which the perturbative cross section is either assumed completely vanishing or at least strongly damped. It suffices to consider hard $2 \to 2$ QCD processes only [52].

Beyond the common basis of the MPI models, they are implemented in different ways by Pythia, Sherpa and Jimmy [53], particularly in the relationship of the additional scatterings with the initial and final state radiation.

In addition, the presence of beam remnants needs to be accounted for. The remnant can be colour-connected to the hard interaction and it forms part of the same fragmenting system.

Many details of the underlying-event modelling involve parameters that have to be tuned using experimental data.

3.2 MC samples

3.2.1 MC samples for the high transverse-momentum top quark analysis

The MC samples used in the high transverse-momentum top quark analysis were simulated using the ATLAS full simulation infrastructure based on Geant4 [54] and reconstructed using the same chain as the data; they are part of the MC11c set of simulations developed to take into account the intricacies of the 2011 data period. The full list of MC samples, their cross-sections and K-factors can be found in Table 3.1. The K factor is defined as $\frac{\sigma^{NNLO}}{\sigma^{MC}}$ and it is used to bring the overall normalisation of any distribution obtained from the MC samples to that of the full NNLO calculation.

The MC signal samples of $t\bar{t}$ events were simulated with the Mc@nlo v4.01 program interfaced to Herwig v6.520 and Jimmy v4.31 for the fragmentation, hadronisation and multi-parton interactions. They use the NLO CT10 [55] parametrisations of the proton PDFs and the AUT2 [56] tune. Only samples in which electrons are part of at least one top-quark decay chain were used. M_{top} was set to 172.5 GeV.

Several sources of background were considered in the analysis. These are W+jets, Z+jets, single-top and dijet QCD processes. Samples of MC events were simulated as follows:

- **W+jets** samples were simulated with the Alpgen v2.13 program interfaced to Herwig v6.520 and Jimmy v4.31. They use the LO CTEQ6L1 [57] parametrisations of the proton PDFs and the AUT2 tune. Only events in which the W boson decays into an electron and a neutrino were used;

- **Z+jets** samples were simulated with the Alpgen v2.13 program interfaced to Herwig v6.520 and Jimmy v4.31. They use the LO CTEQ6L1 parametrisations of the proton PDFs and the AUT2 tune. Only samples in which the Z boson decays into a e^+e^- pair were used;

- **Single top** samples were simulated with the Mc@nlo v 4.01 program interfaced to Herwig v6.520 and Jimmy v4.31, using the NLO CT10 parametrisations of the PDFs

Process	DSID	Number of events	FE*x-sec [pb]	K-factor
ttbar (not all hadronic)	105200	11584773	79.01	1.146
W(enu) + 0 parton	107680	6952874	6930.50	1.196
W(enu) + 1 parton	107681	4998487	1305.3	1.196
W(enu) + 2 parton	107682	3768632	378.13	1.196
W(enu) + 3 parton	107683	1008947	101.86	1.196
W(enu) + 4 parton	107684	744998	25.68	1.196
W(enu) + 5 parton	107685	419947	6.99	1.196
Z(ee) + 0 parton	107650	6618284	668.32	1.25
Z(ee) + 1 parton	107651	1334897	134.46	1.25
Z(ee) + 2 parton	107652	2004195	40.54	1.25
Z(ee) + 3 parton	107653	549949	11.16	1.25
Z(ee) + 4 parton	107654	149948	2.88	1.25
Z(ee) + 5 parton	107655	50000	0.83	1.25
Dijet (J0)	105009	799897	12032000000	
Dijet (J1)	105010	799699	807150000	
Dijet (J2)	105011	799696	48041000	
Dijet (J3)	105012	798995	2536200	
Dijet (J4)	105013	799091	99591	
Dijet (J5)	105014	797693	2594.1	
Dijet (J6)	105015	799284	35.46	
Dijet (J7)	105016	798972	0.13394	
Dijet (J8)	105017	793661	0.0000056775	
Single top(enu) s-channel	108343	253410	0.47	1.064
Single top(enu) t-channel	117360	843473	8.06	0.865
Single top Wt-channel	108346	797024	14.59	1.079

Table 3.1: *Number of events, cross section and K-factor for the different MC samples used in the high transverse momentum top quark analysis.*

and the AUT2 tune, for the s and Wt-channels and with the AcerMC v3.5 program interfaced to Pythia v6.425, using the LO** [58] parametrisations of the proton PDFs and the AUT2B [59] tune, for the t-channel. For the s and t channels only samples in which the W boson decays into an electron and a neutrino were used;

- **QCD dijet samples** were simulated using Pythia v6.425. They use the LO** parametrisations of the proton PDFs and the AUT2B tune.

3.2.2 MC samples for the isolated-photon plus jets analysis

Samples of isolated-photon production were generated by both Pythia 8.165 [60] using the LO CTEQ6L1 parametrisation of the proton PDFs and the AU2 [61] tune and with Sherpa 1.4.0 using the NLO CT10 parametrisation of the proton PDFs. They were simulated using the ATLAS full simulation infrastructure based on Geant4 and reconstructed using the same chain as the data.

The Pythia simulation of the signal includes LO photon plus jets events from both direct processes (the hard subprocesses $qg \to q\gamma$ and $q\bar{q} \to g\gamma$, called hard component) and photon bremsstrahlung in QCD dijet events (called "brem" component). The Sherpa samples were generated with LO matrix elements for photon plus jet final states with up to three additional partons, supplemented with parton showers. The full list of MC samples for the isolated-photon production is shown in Table 3.2. The samples were generated with a minimum E_T^γ requirement.

To avoid overlap it was necessary to cut on the truth E_T^γ of the leading photon in an event to generate exclusive samples and a smooth distribution in E_T^γ.

Process	Run Number	Number of events	FE*x-sec [nb]	E_T Generator [GeV]	E_T range [GeV]
PYTHIA $\gamma + jet$	129172	5949874	1.954e+00	$E_T^\gamma > 70$	$100 < E_T^\gamma < 200$
	129173	1993989	1.184e-01	$E_T^\gamma > 140$	$200 < E_T^\gamma < 400$
	129174	999989	4.843e-03	$E_T^\gamma > 280$	$400 < E_T^\gamma < 650$
	129175	999877	2.088e-04	$E_T^\gamma > 500$	$650 < E_T^\gamma < 1100$
	129176	99997	9.373e-06	$E_T^\gamma > 800$	$1100 < E_T^\gamma$
SHERPA $\gamma + jet$	113715	5499676	2.153e+00	$E_T^\gamma > 70$	$100 < E_T^\gamma < 200$
	113716	2499984	1.379e-01	$E_T^\gamma > 140$	$200 < E_T^\gamma < 350$
	113717	999985	5.963e-03	$E_T^\gamma > 280$	$350 < E_T^\gamma < 650$
	126371	999976	2.765e-04	$E_T^\gamma > 500$	$650 < E_T^\gamma < 1100$
	126955	99996	1.335e-05	$E_T^\gamma > 800$	$1100 < E_T^\gamma$

Table 3.2: *Number of events and cross sections for the different MC samples used in the isolated-photon plus jets analysis.*

Dedicated PYTHIA and SHERPA samples without underlying event were generated at particle and parton levels to correct the NLO calculations for hadronisation and underlying event effects (see section 6.12.1).

Chapter 4

Definition of physics objects

4.1 Electron and photon trigger selection

4.1.1 L1 selection

At L1, trigger information from the EM and hadronic calorimeter system in the form of trigger towers is used. A trigger tower has a dimension of $\Delta\eta \times \Delta\phi$ of 0.1×0.1 in the barrel region. In this region all the cells are summed over the full depth of either the electromagnetic or hadronic calorimeter. The L1 selection algorithm for electromagnetic clusters is based on a sliding 4×4 window of trigger towers which looks for local maxima. An scheme with all the elements that participate in the L1 calorimeter trigger input is shown in figure 4.1.

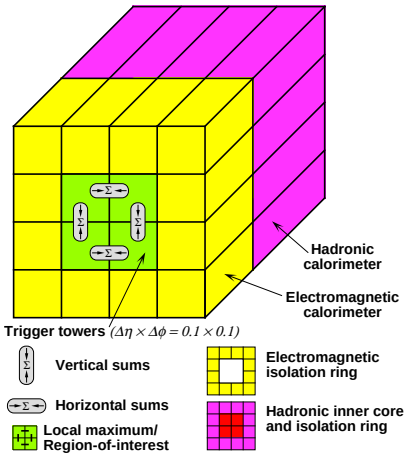


Figure 4.1: *L1 calorimeter trigger scheme, showing how trigger towers are used to determine the energy for the electromagnetic cluster as well as for the electromagnetic isolation, hadronic core and hadronic isolation.*

The trigger object is considered to contain an electron or photon candidate if the following requirements are satisfied:

- the central 2×2 core cluster consisting of both EM and hadronic towers is a local E_T maximum. This requirement prevents double counting of clusters by overlapping windows;
- the most energetic of the four combinations of two neighbouring EM towers passes the electromagnetic cluster threshold.

In addition, it is possible to impose isolation requirements using the 12 EM towers surrounding the core cluster, the 4 hadronic towers behind the core cluster or the 12 hadronic towers

surrounding the core cluster in the hadronic calorimeter.

4.1.2 L2 selection

L2 calorimeter reconstruction is seeded by the η and ϕ positions provided by the L1 trigger. Calorimeter cells in a window of size $\Delta\eta \times \Delta\phi = 0.4 \times 0.4$ are retrieved, which is called Region of Interest (RoI). At the L2 trigger, the cluster building algorithm scans the cells in the second layer of the EM calorimeter and searches for the cell with highest E_T. Subsequently, a cluster of 0.075×0.175 in $\eta \times \phi$ is built around this seed cell in the barrel (0.125×0.125 in the end-cap). Several corrections which are exclusively based on the offline reconstruction algorithms are used at L2 to improve the resolution of the cluster position and energy.

Specifically, the L2 electron and photon algorithm select events based on the following quantities:

- transverse energy of the EM cluster (E_T^{EM}). Due to the energy dependence of the jet cross-section, a cut on E_T^M provides the best rejection against jet background for a given high P_T;
- transverse energy in the first layer of the hadronic calorimeter (E_T^{HAD}): this is required to be below a given threshold. This cut is relaxed for high E_T triggers as the leakage into the hadronic calorimeters increases with energy;
- shower shape in η direction in the second EM sampling: the ratio of the energy deposit in 3×7 cells over that in 7×7 cells is calculated; photons and electrons deposit most of their energy in 3×7 cells and thus the corresponding ratio is typically larger than 80%;
- search for a second maximum in the first EM sampling: the fine granularity in rapidity in the first sampling of the EM calorimeter allows checks to be made for substructures within a shower for a further rejection of background such as single or multiple neutral mesons decaying to photons. The energy deposited in a window $\Delta\eta \times \Delta\phi = 0.125 \times 0.2$ is examined. The shower is scanned for local maxima in the η-direction. The ratio of the difference between the energy deposited in the bin with the highest energy E_{1st} and the energy deposited in the bin with second highest energy E_{2nd} divided by the sum of these two energies is calculated: $R_{strips} = (E_{1st} - E_{2nd})/(E_{1st} + E_{2nd})$. This ratio tends to one for isolated electrons and photons, and tends to zero for photons coming for example from π^0 decay.

If all the criteria of the calorimeter based electron selection are fulfilled, a search for tracks is performed in front of the cluster and electron trigger candidates are identified by the presence of a matching reconstructed track.

The offline track reconstruction is confined to the RoI centred around the L1 seed. The L2 employs a set of algorithms called IDScan for fast pattern recognition and track reconstruction from space points. IDScan first determines de z-position of the primary interaction point along the beam axis and subsequently performs combinatorial tracking only inside groups of space points with the same η and ϕ coordinates and matching the interaction point [62].

4.1.3 EF selection

At the EF trigger level, offline reconstruction algorithms and tools are used as much as possible. The main difference is that the offline reconstruction is run once per event, accessing the whole detector, while the EF uses a seeded approach, i.e. it runs several times per event, once for each RoI given by L2, accessing only the corresponding subsample of the detector.

EM clusters are searched for and reconstructed in RoIs of size $\Delta\eta \times \Delta\phi = 0.4 \times 0.4$. The algorithms look for a local energy maximum with calorimeter trigger tower granularity. For electron and photon reconstruction only the data from the EM calorimeter is used, in contrast with L2 where the hadronic energy is also computed. The clusters should have an E_T above a given threshold. The default cluster size is 0.125×0.125 in $\eta \times \phi$. Once found by the clustering algorithm, the cluster parameters (position, energy, etc.) are computed and further refined by a set of cluster correction tools.

For electron triggers, tracks are reconstructed in the inner detector. The EF tracking implements an adapted version of the offline reconstruction software, constrained to the data available in the RoI.

Electron and photon identification in the EF is very similar to the offline one, which is described in sections 4.3 and 4.4. Calorimeter shower shapes, leakage into the hadronic calorimeter and the E_T of the EM cluster are used for the calorimeter based selection of electrons and photons. For electrons track-cluster matching variables, track quality cuts and transverse impact parameters are also used.

An study on the efficiency of the photon trigger used in this dissertation is included in section 7.2.

4.2 Jet trigger selection

Jets at the trigger level are reconstructed from Jet Elements (JE), formed from the energy sum of 2×2 trigger towers in both the EM and the hadronic calorimeters. JEs have a basic granularity of 0.2×0.2 in $\Delta\eta \times \Delta\phi$. L1 jet candidates are identified using a sliding window of JE energy sums in either a 2×2, 3×3 or 4×4 JE region, depending on the trigger element. A jet is reconstructed if the total transverse (EM + Hadronic) energy within the window is above a given threshold. To prevent the L1 algorithm from identifying overlapping jets, the transverse energy of a cluster, defined as a region spanned by 2×2 JE, is required to be a local maximum within ± 0.4 units in both η and ϕ. Figure 4.2 shows a schematic diagram of the jet reconstruction algorithm at L1.

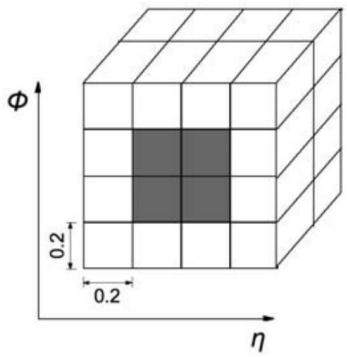


Figure 4.2: *Schematic diagram of the L1 jet algorithm showing a window of* 4×4 *JEs spanning the electromagnetic and hadronic calorimeter in depth, and a local maximum transverse energy cluster of* 2×2 *JEs.*

In contrast to other calorimeters, the L1 trigger towers in the FCAL have a granularity of 0.4×0.4 in $\Delta\eta \times \Delta\phi$. As a consequence, a JE in the FCAL is formed by summing calorimeter towers in η. The FCAL JEs have a ϕ granularity of 0.4 with only a single η bin at each end.

For each selected event, the L1 calorimeter subsystems provide RoI information, where jet candidates have been identified, to the L2 system. Using the RoI locations and calorimeter cells

within RoIs, the L2 jet trigger uses a simple three iteration cone Feature Extraction (FEX) algorithm to build jets [62]. Only one jet can be produced per RoI. A hypothesis algorithm then tests to determine if the located jet meets the trigger selection criteria, usually just a minimum E_T requirement.

The final stage of the trigger, the EF, uses the standard ATLAS analysis and event reconstruction applications developed for offline analysis, as well as the final calibration. In addition, the EF is able to access the information on the entire event, rather than just within RoIs. The EF jet trigger reconstructs jets with the anti-k_{T} algorithm over the entire detector in a single pass in an identical manner to the standard offline jet reconstruction, which is described in detail in section 4.5.

4.3 Electrons in ATLAS

The definitions and procedures included in this section are described as they were implemented for the high transverse-momentum top quark analysis included in this dissertation and follow the guidelines recommended by the ATLAS collaboration for analysis of the 2011 $\sqrt{s} = 7$ TeV dataset at the moment of its completion.

4.3.1 Electron reconstruction

The electron reconstruction algorithm used in the central region of the detector, equipped with the ID ($|\eta| < 2.5$), identifies energy deposits in the EM calorimeter and associates these clusters of energy with reconstructed tracks in the inner detector. The three step process is as follows:

- **cluster reconstruction:** EM clusters are seeded from energy deposits with total transverse energy above 2.5 GeV by using a sliding-window algorithm with window size 3×5 in units of 0.025×0.025 in $\Delta\eta \times \Delta\phi$ space. From MC simulations of W and Z leptonic decays, the efficiency of the initial cluster reconstruction is expected to be approximately 97% at $E_T = 7$ GeV and almost 100% for electrons with $E_T > 20$ GeV [63];

- **track association with the cluster:** within the tracking volume, tracks with $P_T > 0.5$ GeV are extrapolated from their last measured point to the middle layer of the EM calorimeter. The extrapolated η and ϕ coordinates of the impact point are compared to a corresponding seed cluster position in that layer. A track and a cluster are considered to be successfully matched if the distance between the track impact point and the EM cluster barycentre is $|\Delta\eta| < 0.05$. To account for the effect of bremsstrahlung losses on the azimuthal distance, the size of the $\Delta\phi$ track-cluster matching window is 0.1 on the side where the extrapolated track bends as it traverses the solenoidal magnetic field. An electron candidate is considered to be reconstructed if at least one track is matched to the seed cluster. In the case where more than one track is matched to a cluster, tracks with hits in the pixel detector or the SCT are given priority and only the closest track to the cluster is used. In the absence of a matching track, the cluster is classified as an unconverted photon candidate. Electrons are distinguished from converted photons by investigating the presence of pairs of close-by tracks originating from a vertex displaced from the interaction point and by verifying the location of the first hits along the path of the single tracks [64];

- **reconstructed electron candidate:** after a successful track-cluster matching, the cluster sizes are optimised to take into account the overall energy distributions in the different regions of the calorimeter. In the EM barrel region, the energy of the electron cluster is

collected by enlarging its size to 3×7 in units of 0.025×0.025 in $\Delta\eta \times \Delta\phi$, while in the EM end-caps the size is increased to 5×5. The $\eta - \phi$ spatial coordinates of the electron candidates are taken from the parameters of the matched track at the interaction vertex. The cluster energy is then determined by applying correction factors computed by a calibration scheme based on the full detector simulation, which is described in section 4.3.4. Electrons reconstructed near regions of the calorimeter affected by read-out failures are rejected [65].

4.3.2 Electron identification

The identification criteria for electron candidates are implemented based on sequential cuts on calorimeter, tracking and combined track-cluster variables. These requirements are optimised in 10 cluster-η bins, motivated by the structure of the detector, and 11 E_T bins (from 5 to 80 GeV), in order to provide good separation between signal electrons and background from hadrons misidentified as electrons or electrons from photon conversions.

Three sets of reference selection criteria, labelled *loose*, *medium* and *tight*, are designed to identify electrons. These criteria are designed in a hierarchical way so as to provide increasing background rejection power at some cost of identification efficiency. The different selections used for central-electron identification are detailed below:

- *loose:* the loose selection uses shower-shape variables in both the first and second layers of the EM calorimeter. Hadronic leakage information is also used. Additional requirements on the quality of the electron track and track-cluster matching help to improve rejection of hadronic backgrounds, particularly in the 30 to 40 GeV E_T range:
  - R_{had}: ratio of the E_T in the first layer of the hadronic calorimeter to the transverse energy of the EM cluster. In the range $0.8 < |\eta| < 1.37$, which is not covered by the first hadronic layer, the total hadronic transverse energy to the EM transverse energy ratio is used.
  - middle layer of the EM:
    - R_η: ratio of energies in 3×7 cells over 7×7 cells;
    - w_2: lateral width of the shower.
  - front layer of the EM:
    - w_{stot}: total shower width;
    - E_{ratio}: energy difference between the largest and second largest energy deposits in the cluster divided by their sum.
  - number of hits in the pixel detector (> 0);
  - number of hits in the silicon detectors (≥ 7).
  - $\Delta_{\eta 1}$: $|\Delta\eta|$ between the cluster position in the first layer and the extrapolated track (< 0.015).
- *medium:* the medium selection modifies the loose one by requiring the presence of a measured hit in the innermost layer of the pixel detector (to reject electrons from photon conversions); applying a loose selection requirement on the transverse impact parameter and identifying transition radiation in the TRT (to reject charged-hadron background) when available. The requirements on the discriminating variables in common with the loose selection are also tightened, allowing the background rejection power to increase by approximately an order of magnitude with respect to the loose selection:
  - number of hits in the innermost layer of the pixel detector (> 0) for $|\eta| < 2.01$;

  - number of hits in the pixel detector (> 1) for $|\eta| > 2.01$;
  - d_0: transverse impact parameter $|d_0| < 5mm$;
  - loose cut on TRT high-threshold fraction.

- *tight:* the tight selection makes full use of the particle-identification tools available for electron identification. In addition to the generally tighter requirements on medium selection discriminating variables, stricter requirements on track quality in the presence of a track extension in the TRT detector, on the ratio of the EM cluster energy to the track momentum and a veto on reconstructed photon conversion vertices associated with the cluster are applied. Overall, a rejection power higher by a factor of two is achieved with respect to the medium selection:
  - tighter transverse impact parameter cut $|d_0| < 1$mm;
  - asymmetric cut on $\Delta\phi$ between the cluster position in the middle layer and the extrapolated track;
  - E/p: ratio of the cluster energy to the track momentum;
  - total number of hits in the TRT;
  - tighter cut on the TRT high-threshold fraction;
  - rejection of electron candidates matched to reconstructed photon conversions.

It is important to note that none of the electron identification criteria explicitly apply requirements on the presence of other particles (additional tracks or energy deposits outside the EM cluster) close to the identified electrons. The optimisation of such requirements, so-called isolation requirements, is treated separately in section 4.3.3. In Figure 4.3 the distributions, obtained from MC simulations, for some of the variables used in electron identification are shown. The four categories shown in the plots are defined based on a match to MC particle level as:

- isolated electrons: if they match a true electron originating from a Z or W boson;
- hadron fakes: if they do not match a true electron, muon or tau;
- non-isolated electrons: if they match a true electron originating from b-mesons;
- background electrons: if they match a true electron coming from Dalitz decays or coming from a photon.

4.3.3 Electron isolation

Various analyses may require different isolation criteria of different tightness and therefore the isolation cuts are applied on top of the electron identification. Different calorimeter and tracking based isolation variables have been studied and optimised by the ATLAS collaboration.

A calorimetric isolation discriminator (E_T^{iso}) is computed from the reconstructed energy in a cone of half opening angle R_0 around the electron candidate direction, where the energy of the electron itself is excluded. Cells from both the EM and hadronic calorimeters are included. While a larger cone will contain more energy in case of misidentified jets, a smaller cone is more robust against pile-up effects. For this dissertation, a value of $R_0 = 0.2$ was used.

In addition, a tracking based discriminator is used, which is the scalar sum of the P_T of tracks in a cone of $R_0 = 0.3$ around the electron. In contrast to the calorimetric isolation, neutral particles do not contribute to this quantity. The advantage, however, is that track quality criteria can be applied in order to reject tracks from secondary vertices. Track isolation

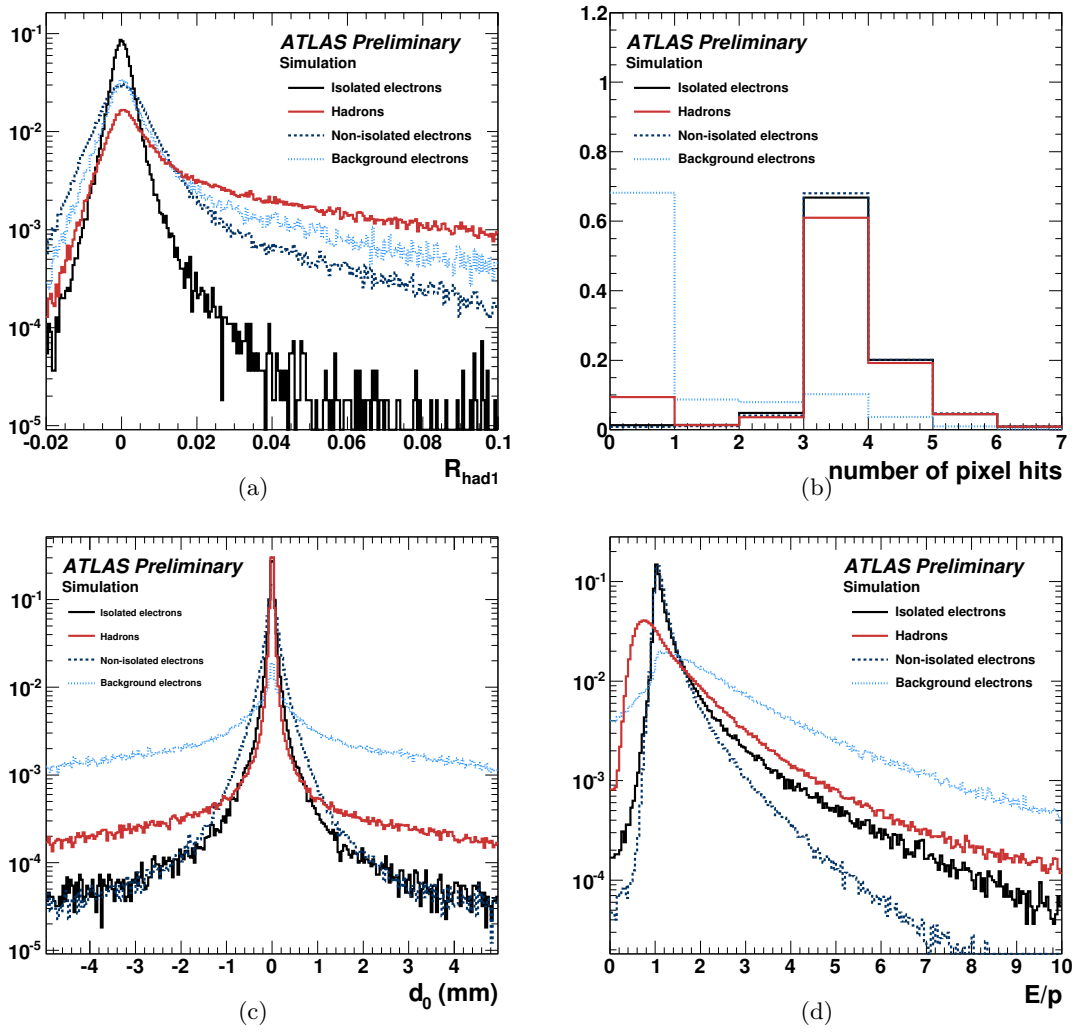


Figure 4.3: *R_{had1} (a), number of pixel hits (b), impact parameter (c) and E/p distributions, obtained from MC simulations, for isolated electrons and the main backgrounds to isolated electron studies [66].*

considers tracks with $P_T > 1$ GeV, a hit in the innermost pixel detector layer, at least 7 hits in silicon detectors and transverse and longitudinal impact parameter less than 1mm. For tracks with $\Delta R < 0.1$ with respect to the electron it is also required that they are not matched to a conversion vertex.

Isolation cuts are optimised for different regions in E_T and $|\eta|$ and it is done in such a way that the efficiency for isolated electrons is constant in all ranges of P_T and $|\eta|$. Figure 4.4 shows the distributions of both isolation discriminators (normalised) for isolated electrons coming from $Z \to ee$ decays and hadrons faking electrons. They were obtained using a PYTHIA $Z \to ee$ sample for the signal and a PYTHIA di-jet sample for the background.

The base calorimetric energy described in this section is affected in two unwanted ways and thus dedicated corrections are applied:

- an electron will leak some of its energy outside of its central core, and will cause the isolation energy to grow as a function of the electron E_T:
  - **leakage corrections** are obtained from single particle MC samples and subtracted from E_T^{iso}.
- soft energy deposits from interactions different from the hard scattering generating the

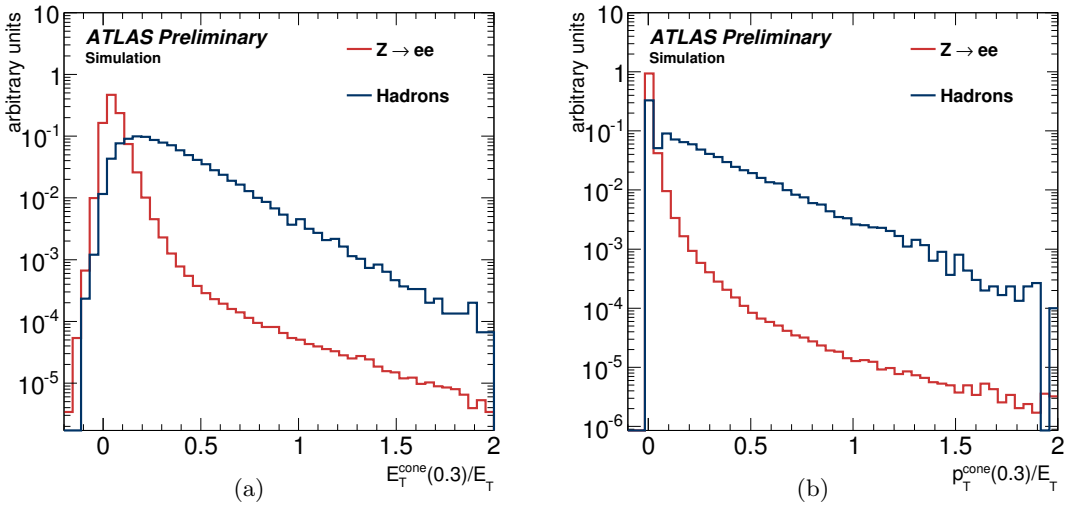


Figure 4.4: *Calorimetric and tracking isolation energy normalised to the electron transverse energy for a cone of $R_0 = 0.3$ The electron candidates are required to have $E_T > 15$ GeV and to pass the medium electron identification. The samples used are a* Pythia *$Z \to ee$ sample for the signal and a* Pythia *di-jet sample for the background [66].*

electron will contribute to the isolation energy depending on the amount of activity in the current event, both from the underlying event and "in-time pile-up", as well as activity coming from previous events (out-of-time pile-up):

– **Pileup corrections** are parametrised as functions of the number of primary vertices (N_{pv}) and subtracted from E_T^{iso}. The corrections used for this dissertation were obtained using the full 2011 $\sqrt{s} = 7$ TeV dataset.

4.3.4 Electron energy scale calibration

The strategy to calibrate the electron energy is divided into two steps:

- a MC-based calibration applied at the cluster level for energy loss due to absorption in the passive material and leakage outside the cluster. It includes additional fine corrections depending on the η and ϕ coordinates of the electron which are made to compensate for the energy modulation as a function of the impact point in the calorimeter;
- in-situ calibration using $Z \to ee$ decays determines the energy scale and inter-calibrates the different regions of the calorimeters.

MC-based calibration

In this calibration procedure, known as the calibration-hits-method [67], special simulations are used in which the energy deposited by a particle is recorded in all detector materials, not just the active ones. Through these simulations, the energy depositions in the inactive material can be correlated with the measured quantities to parametrised them and make it possible to estimate them when reconstructing electrons.

The cluster energy is determined precisely by computing and summing four different contributions: the energy deposited in the material in front of the EM calorimeter, that deposited in the calorimeter inside the cluster, that deposited outside the cluster (lateral leakage) and the energy deposited beyond the EM calorimeter (longitudinal leakage). The four terms are

parametrised as a function of the measured energies in the pre-sampler (PS), where it is available, and in the three accordion longitudinal layers. The parameters are computed at each pseudorapidity value corresponding to the centre of a cell in the middle layer. Symmetry in pseudorapidity is assumed.

More precisely, the energy of the produced electron is reconstructed with the following formula

$$E_e = \underbrace{a(E_{tot}^{Acc},\eta) + b(E_{tot}^{Acc},\eta)\cdot E_{ps}^{clLAr} + c(E_{tot}^{Acc})\cdot (E_{ps}^{clLAr})^2}_{\text{Energy in front}} + \underbrace{\frac{s_{cl}^{Acc}(X,\eta)}{f_{out}(X,\eta)}\cdot(\sum_{i=1,3} E_i^{clLAr})}_{\text{Energy in the accordion}}\cdot \underbrace{(1+f_{leak})}_{\text{Longitudinal leakage}} \cdot \underbrace{F(\eta,\phi)}_{\text{Energy modulation}} , \tag{4.1}$$

where:

- E_e is the electron energy;
- $a(E_{tot}^{Acc},\eta)$, $b(E_{tot}^{Acc},\eta)$ and $c(E_{tot}^{Acc},\eta)$ are parameters determined as a function of the energy deposited in the accordion (E_{tot}^{Acc}) and $|\eta|$;
- E_{ps}^{clLAr} is the part of the cluster energy measured in the pre-sampler, corrected for the fraction deposited in the passive materials;
- X is the longitudinal barycentre of the shower defined by

$$X = \frac{\sum_{i=0}^{3} E_i^{clLAr}\cdot X_i}{\sum_{i=0}^{3} E_i^{clLAr}} \tag{4.2}$$

  ,

  where E_i^{clLAr} are the energies deposited in the cluster in the active medium of the pre-sampler and the three compartments of the calorimeter (strip, middle and back) and X_i is the depth, expressed in radiation lengths, of the longitudinal centre of each compartment computed from the centre of ATLAS. The X_i depends on $|\eta|$;
- $s_{cl}^{Acc}(X,\eta)$ is a correction factor to the accordion sampling fraction in the cluster;
- $f_{out}(X,\eta)$ is the correction for the energy deposited in the calorimeter outside the cluster;
- $f_{leak}(X,\eta)$ is the longitudinal leakage correction;
- $F(\eta,\phi)$ is the energy correction depending on the impact point inside a cell (energy modulation).

In the region $|\eta| \geq 1.8$, not instrumented with the pre-sampler, the energy deposited in front of the calorimeter is parametrised as a function of X computed only with the information given by the three calorimeter layers.

The linearity and resolution of the calibration hits based method is shown for electrons in figure 4.5.

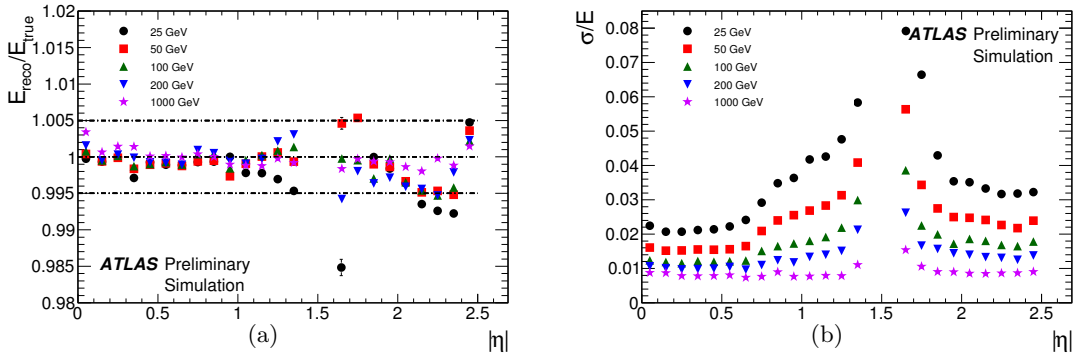


Figure 4.5: *Linearity (a) and resolution (b) obtained by the calibration hits method as a function of $|\eta|$ for different electron energies [68].*

Energy-scale determination using dielectron Z decays

Any residual miscalibration for a given region i in $|\eta|$ is parametrised as

$$E^{meas} = E^{true}(1 + \alpha_i), \tag{4.3}$$

where E^{true} is the true electron energy, E^{meas} is the energy measured by the calorimeter after the MC-based energy-scale correction, and α_i measures the residual miscalibration. The α_i that correct the energy scale are determined by a fit minimising the negative unbined log-likelihood

$$-\ln L_{tot} = \sum_{i,j} \sum_{k=1}^{N_{ij}^{events}} (-1) \ln L_{ij}\left(\frac{m_k}{1 + \frac{\alpha_i + \alpha j}{2}}\right), \tag{4.4}$$

where the indices i, j denote the regions considered for the calibration with one of the electrons from the $Z \to ee$ decay being in region i and the other in region j, N_{ij}^{events} is the total number of selected $Z \to ee$ decays with electrons in regions i and j, m_k is the measured dielectron invariant mass in a given decay and $L_{ij}(m)$ is the probability density function (pdf) quantifying the compatibility of an event with the Z line shape. This pdf template is obtained from PYTHIA MC simulations and smoothed to get a continuous distribution. The resulting α values obtained by applying the procedure described above to the full 2010 $\sqrt{s} = 7$ TeV dataset in 58 η bins are shown in figure 4.6. The corrections applied for this dissertation are based on the 2011 $\sqrt{s} = 8$ TeV dataset.

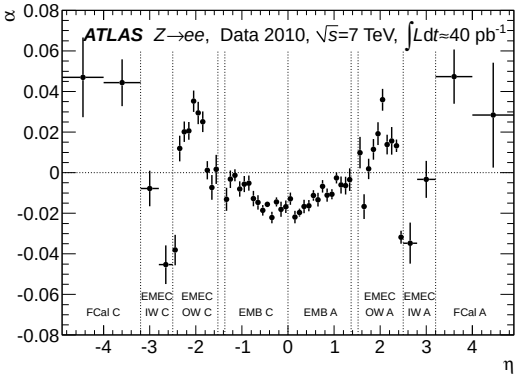


Figure 4.6: *Energy scale correction factor as a function of η derived from fits to $Z \to ee$ events using 2010 data [69].*

4.4 Photons in ATLAS

The definitions and procedures included in this section are described as they were implemented for the isolated-photon plus jets analysis included in this dissertation and follow the guidelines recommended by the ATLAS collaboration for analysis of the 2012 $\sqrt{s} = 8$ TeV dataset.

4.4.1 Photon reconstruction

Photon reconstruction starts with exactly the same calorimeter based procedure that the electron reconstruction described in section 4.3.1. Clusters with no associated tracks are classified as unconverted photons.

Converted photons, photons that have converted to a e^+e^- pair, are recovered from electron candidates if the matched track is consistent with originating from a photon conversion and if, in addition, a conversion vertex is reconstructed. They are classified as single-track or double-track conversions depending on the number of assigned electron-tracks [70].

Converted photons are reconstructed using a 3×7 cluster size in $\Delta\eta \times \Delta\phi$ in the EM barrel while unconverted photons are reconstructed using a 3×5 cluster due to their smaller lateral size. A 5×5 cluster size is used in the EMEC for converted and unconverted photons.

Photons reconstructed near regions of the calorimeter affected by read-out failures are rejected [65].

4.4.2 Photon identification

The *tight* photon identification criteria rejects hadronic background by applying requirements on several variables related to the energy leaking into the hadronic calorimeter and the shower shape.

The nine discriminating variables used in this identification are:

- R_{had}: ratio of the E_T in the first layer of the hadronic calorimeter to the transverse energy of the EM cluster. In the range $0.8 < |\eta| < 1.37$, which is not covered by the first hadronic layer, the total hadronic transverse energy to the EM transverse energy ratio is used;
- middle layer of the EM:
  - R_η: ratio of energies in 3×7 cells over 7×7 cells;
  - w_2: lateral width of the shower;
  - R_ϕ: ratio of energies in 3×3 cells over 3×7 cells.
- front layer of the EM:
  - w_{s3}: the shower width for three strips around the maximum strip;
  - w_{stot}: total shower width;
  - F_{side}: the fraction of energy outside a core of 3 central strips, but within 7 strips;
  - ΔE: the difference between the energy of the strip with the second-largest energy deposited and the energy of the strip with the smallest energy deposit between the two leading strips;
  - E_{ratio}: energy difference of the largest and second largest energy deposits in the cluster divided by their sum.

The selection based on R_{had} and the middle layer variables aims to reject most of the jet background. The selection on the front layer variables is optimised to reduce further the contribution from jets with one or more hard π^0's decaying to photons and carrying most of the jet energy, which would pass the previous requirements.

Three additional looser selections for photons were used in this dissertation. They are obtained by imposing requirements on subsamples of the same variables that are used in the *tight* selection:

- *loose prime*: based only on R_{had}, R_η, w_2, R_ϕ and w_{stot};
- *tighter loose prime*: adding ΔE and w_{s3} to the loose prime requirements;
- *looser loose prime*: removing w_{stot} from the loose prime requirements.

In figure 4.7 the distributions for the 9 tight variables, obtained from MC simulations, are shown for true and fake electrons classified as unconverted before any selection.

4.4.3 Photon isolation

Photon isolation requirements are based on the isolation transverse energy, E_T^{iso}. The E_T^{iso} is reconstructed by using topo-clusters (see section 4.5.1) calibrated at the EM scale within a cone of radius $R = 0.4$ in the $\eta - \phi$ plane around the photon cluster barycentre. Only positive energy topo-clusters are used. The topo-clusters include cells from the EM and hadronic calorimeter, but the TileGap3 cells are explicitly removed. The energy from the core of the cone in the electromagnetic calorimeter is subtracted from E_T^{iso}.

Leakage corrections are evaluated as a function of E_T^γ on simulated samples of single photons and subtracted from E_T^{iso}.

Further corrections are also applied to E_T^{iso} to remove the effects from the underlying event and pile-up and to match the definition between data and theory. This correction is obtained by the so-called *jet-area* method [71]. In this method, low-energy jets are used to compute the ambient energy density in an event-by-event basis, which is then multiplied by the area of the isolation cone and subtracted from the isolation energy. The ambient energy per unit of area reconstructed in the 2012 data-taking period is shown in figure 4.8. The correction is typically between 1.5 and 2 GeV.

A dedicated analysis of the photon-plus jets analysis isolation cuts is discussed in section 7.1.

4.4.4 Photon energy scale calibration

The photon energy calibration proceeds as a three step process:

- the photon response in the EM calorimeters is calibrated using a Multivariate algorithm (MVA) based calibration;
- the response of the LAr layer scales (PS and Middle Layer E2) are corrected in data before applying the MVA-based calibration. After this initial pre-calibration, additional corrections accounting for effects related to the LAr gain switch miscalibration, the LAr intermodule widening and the imperfect LAr HV uniformity are applied;
- the final step of the calibration procedure consists on applying in-situ energy scales obtained from the comparison of the detector response to $Z \to ee$ events in data and MC, which are assumed to be valid also for photons.

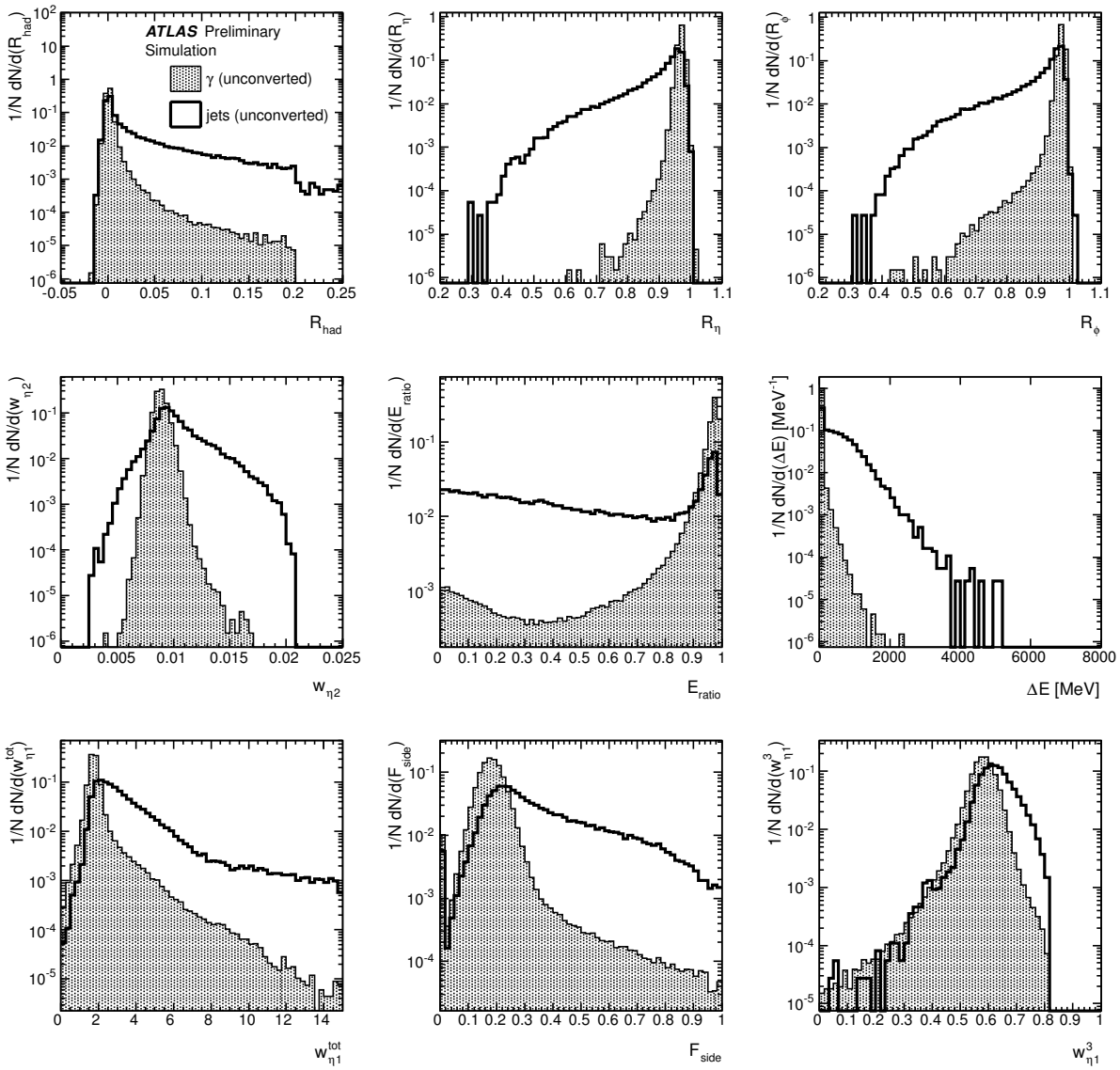


Figure 4.7: *Normalised distributions, obtained from MC simulation, of the calorimetric discriminating variables in the region* $0 < |\eta| < 0.6$ *for* $E_T^\gamma > 20$ *GeV for true and fake photons reconstructed as unconverted before any selection [64].*

MVA based correction

Reconstructed photons clusters are calibrated to correct for the energy lost in the material upstream of the calorimeter, the energy deposited in the cells neighbouring the cluster in η and ϕ and the energy lost beyond the LAr calorimeter. Further corrections are also applied to correct for the response dependence as a function of the particle impact point within the central cluster cell.

The corrections are determined using a multivariate algorithm and are applied separately for converted and unconverted photons in η and p_{T} bins. The calibration procedure optimises the estimate of the true particle energy (E_{true}) from the detector-level observables.

The quantities used for electrons and photons are the total energy measured in the calorimeter E_{calo}, the ratio of the PS energy to the calorimeter energy, the longitudinal barycentre defined in equation 4.2, the cluster barycentre pseudorapidity in the ATLAS coordinate system ($\eta_{cluster}$) and the cluster barycentre in η and ϕ within the calorimeter frame.

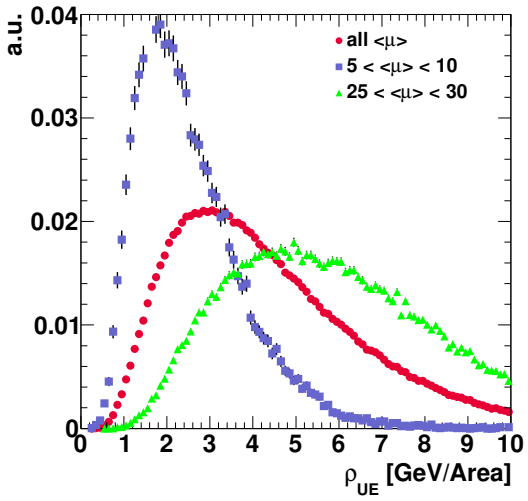
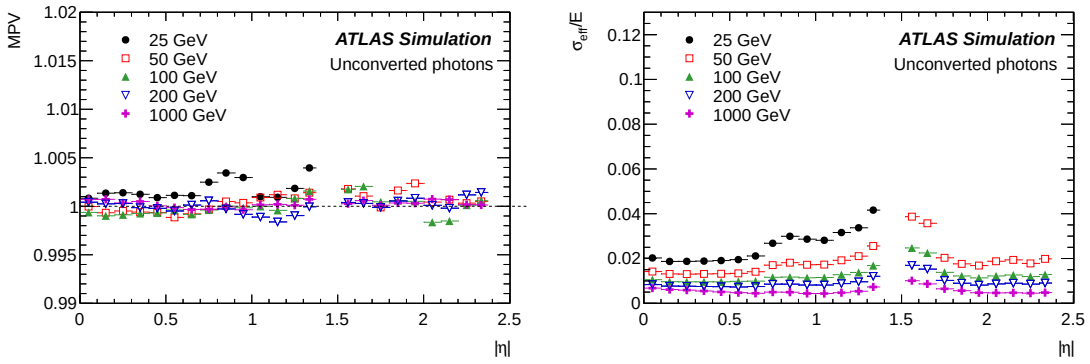


Figure 4.8: *Ambient energy per unit of area reconstructed during 2012 data taking.*

The variable $\eta_{cluster}$ is included to account for the passive-material variations in front of the calorimeter. The inclusion of the barycentre location in the calorimeter frame is important to accurately correct for the increase of lateral energy leakage of photons that hit the cell close to the edge.

For converted photons, the conversion radius R_{conv} is also used as an additional input to the MVA.

The linearity and resolution of the MVA-based calibration method is shown for unconverted photons in figure 4.9.



(a) (b)

Figure 4.9: *Most probable value of E/E_{true} (a) and resolution (b) for the MVA-based calibration method as a function of η for different energies of unconverted photons [70].*

Energy-scale determination using dielectron decays of Z

As already explained for electrons, any residual miscalibration can be parametrised using equation 4.3.

For photons in analyses performed with the 2012 $\sqrt{s} = 8$ TeV dataset, the extraction of α_i is performed differently as the one described in section 4.3.4. The α_i are obtained simultaneously as the resolution corrections, which are derived under the assumption that the resolution curve

is well modelled by the simulation up to a Gaussian constant term:

$$\left(\frac{\sigma_E}{E}\right)^{data} = \left(\frac{\sigma_E}{E}\right)^{MC} \oplus c \tag{4.5}$$

For $Z \to ee$ events, it is possible to write the following equations:

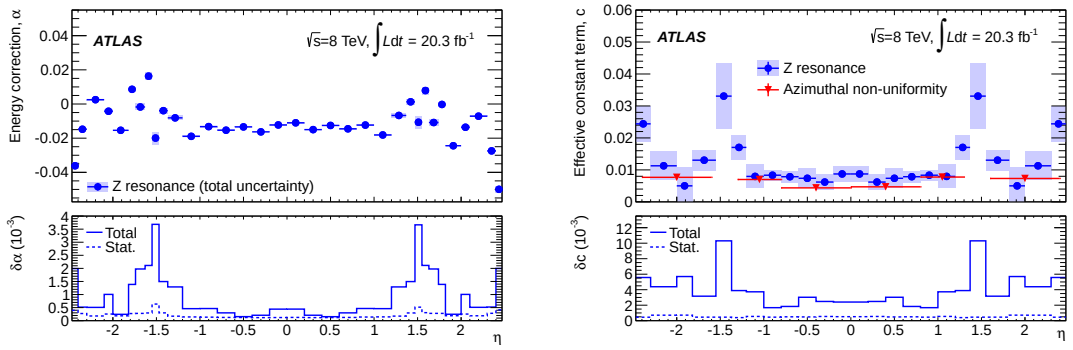
$$m_{ij}^{data} = m_{ij}^{MC}(1 + \alpha_{ij}), \tag{4.6}$$

and

$$\left(\frac{\sigma_m}{m}\right)_{ij}^{data} = \left(\frac{\sigma_m}{m}\right)_{ij}^{MC} \oplus c_{ij} \tag{4.7}$$

where $\alpha_{ij} = (\alpha_i + \alpha_j)/2$ and $c_{ij} = c_i \oplus c_j/2$. For each (η_i, η_j) category, templates are built with energy scale and resolution perturbations to the detector-level quantities, in a range covering the expected uncertainty in narrow steps, building a two dimensional grid along (α_{ij}, c_{ij}). The optimal values, uncertainties and correlations for α_{ij} and c_{ij} are obtained by χ^2 minimisation.

The energy scale corrections and the constant term corrections obtained with this method for the 2012 $\sqrt{s} = 8$ TeV analysis are shown in figure 4.10.



(a) (b)

Figure 4.10: *Energy scale corrections (a) and effective constant term corrections (b) as a function of η, derived from $Z \to ee$ events using the template method. The bottom plots shows the statistical and total uncertainties. The values of c are symmetrised with respect to $\eta = 0$ [70].*

4.5 Jets in ATLAS

Jets were used in both analyses included in this dissertation. In this section and for simplicity, *2011 jets* refer to the jets as they were defined for the high transverse-momentum top quark analysis performed using the 2011 $\sqrt{s} = 7$ TeV and *2012 jets* refer to the jets as they were defined for the isolated-photon plus jets analysis performed using the 2012 $\sqrt{s} = 8$ TeV.

Jets were reconstructed using three different inputs:

- jets at the detector-level, reconstructed using topological clusters in the calorimeter. Same definition is used for data and MC samples;
- jets at the particle-level, reconstructed using stable particles (lifetime longer than 10 ps) from MC simulations. Muons and neutrinos were not included;

- jets at the parton-level, reconstructed using partons from NLO calculations or MC simulations.

The two algorithms described in section 1.3, k_{T} and anti-k_{T} were used with several radius parameters, always utilising the FastJet software package [72]

The discussion in this section focuses on the first type of jets, those at the detector-level.

4.5.1 Calorimeter jet reconstruction

The input to calorimeter jets are topological clusters of calorimeter cells, called topo-clusters, constructed from adjacent calorimeter cells that contain a significant energy signal above noise [73].

Topo-cluster reconstruction

The clustering starts with a seed cell, whose signal-to-noise ratio is above a threshold of $S/N = 4$. The noise is estimated using:

$$\sigma_{noise} = \sqrt{(\sigma_{noise}^{electronic})^2 + (\sigma_{noise}^{pile-up})^2} \tag{4.8}$$

where $\sigma_{noise}^{electronic}$ is estimated as the absolute value of the energy deposited in the calorimeter cell divided by the RMS of the energy distribution measured in events triggered at random bunch crossings and $\sigma_{noise}^{pile-up}$ is determined from MC simulations [74]. Cells neighbouring the seed (or the cluster being formed) that have a signal-to-noise ratio of at least $S/N = 2$ are included iteratively. Finally, all calorimeter cells neighbouring the formed topo-cluster are added. The topo-cluster algorithm efficiently suppresses the calorimeter noise.

The topo-cluster algorithm also includes a splitting step in order to optimise the separation of showers from different close-by particles. All cells in a topo-cluster are searched for local maxima in terms of energy content with a threshold of 500 MeV. This means that the selected calorimeter cells have to be more energetic than any of its neighbours. The local maxima are then used as seeds for a new iteration of topological clustering, which splits the original clusters into more topo-clusters.

Finally, the topo-cluster is defined to have an energy equal to the energy sum of all the included calorimeter cells, zero mass and a reconstructed direction calculated from the weighted averages of the pseudorapidities and azimuthal angles of the constituent cells. The weight used is the absolute cell energy and the positions of the cells are relative to the nominal ATLAS coordinate system. This procedure yields topo-clusters at the electromagnetic scale (EM-scale) [75], which correctly measures the energy deposited in the calorimeter by particles produced in electromagnetic showers.

A second topo-cluster collection is built by taking the previous EM-scale collection and calibrating the calorimeter cells in these clusters such that the response of the calorimeter to hadrons is correctly accounted for. This calibration uses the local cluster weighting method that aims at an improved resolution compared to the EM scale by correcting for a variety of effects in the calorimeter:

- clusters are classified as electromagnetic or hadronic such that the non-compensating nature of the ATLAS calorimeter can be accounted for;
- the energy falling outside clustered cells is estimated from how isolated the cluster is;
- the amount of energy falling in inactive areas of the detector is estimated from the position and energy deposited in each layer of the calorimeter.

The corrections are obtained from simulations of charged and neutral particles. The topo-clusters are calibrated *locally*, without considering the jet context. This procedure is known as local cluster calibration (LCW).

4.5.2 Jet quality

Jets with high transverse momentum must be distinguished from background jet candidates not originating from hard-scattering events. The main sources of potential background are:

- beam-gas events, where one proton of the beam collides with the residual gas within the beam pipe;
- beam-halo events, for example caused by interactions in the tertiary collimators in the beam-line far away from the ATLAS detector;
- cosmic-ray muons overlapping in-time with collision events;
- calorimeter noise.

Beam induced background (BIB) and cosmic rays

The distribution of energy deposits by the jet, the shower shape and its direction, in particular the pointing to the interaction point, can be employed to discriminate collision jets from beam induced fake jets. Examples of discriminating variables are the electromagnetic energy fraction (f_{em}), defined as the energy deposited in the electromagnetic calorimeter, divided by the total jet energy and the maximum energy fraction in any single calorimeter layer (f_{max}).

The vast majority of collision jets contain charged hadrons that are reconstructed by the tracking system. In the tracker acceptance, $|\eta| < 2.5$, the jet charged particle fraction (f_{ch}) is defined as the ratio of the scalar sum of the P_T of the tracks associated with the jet divided by P_T^{jet}. This is a powerful tool to discriminate collision jets from fake jets, which typically have no associated tracks. Finally, beam induced and cosmic rays induce jet candidates that are usually not in-time with the collision products.

Noise in the calorimeters

Most of the noise is already identified and rejected by the data quality inspection performed shortly after data-taking, based on standardised quality criteria. A small fraction of calorimeter noise remains undetected and needs to be rejected by additional criteria. The characteristic pulse shape of real energy deposits in the calorimeter cells can be used to distinguish a true ionisation signal from noise. This leads to the definition of quality variables:

- f_{HEC}: fraction of the jet energy in the HEC calorimeter;
- $\langle Q \rangle$: the average jet quality is defined as the energy-squared weighted average of the pulse quality of the calorimeter cells $\langle Q_{cell} \rangle$ in the jet. This quantity is normalised such that $0 < \langle Q \rangle < 1$;
- F_Q^{LAr}: fraction of the energy in LAr calorimeter cells with poor signal shape quality ($Q_{CELL} > 4000$);
- F_Q^{HEC}: fraction of the energy in HEC calorimeter cells with poor signal shape quality ($Q_{CELL} > 4000$);

- E_{neg}: energy of the jet originating from cells with negative energy that can arise from electronic noise of early out-of-time pile-up or from sporadic noise burst.

In figure 4.11, distributions for collision and fake jets reconstructed using the 2011 $\sqrt{s} = 7$ TeV dataset for several of the quality variables are shown after having passed four successive sets of jet quality criteria, *Looser*, *Loose*, *Medium* and *Tight* defined using the variables described above. The looser quality criteria are designed to provide a signal efficiency above 99.8% with a fake-jet rejection factor of about 50% while the tight criteria are designed to provide a large fake-jet rejection factor with a signal inefficiency not larger than a few percent. The 4 quality criteria used to identify fake jets are defined as:

- ***looser* criteria:**
  - **BIB and cosmic rays:** ($f_{max} > 0.99$ and $|\eta| < 2.0$) or ($f_{em} < 0.05$ and $f_{ch} < 0.05$ and $|\eta| < 2.0$) or ($f_{em} < 0.05$ and $|\eta| \geq 2.0$);
  - **calorimeter noise:** ($f_{hec} > 0.5$ and $|f_Q^{hec}| > 0.5$ and $\langle Q \rangle > 0.8$) or ($E_{neg} > 60$ GeV) or ($f_{em} > 0.95$ and $f_Q^{LAr} > 0.8$ and $\langle Q \rangle > 0.8$ and $|\eta| < 2.8$)
- ***loose* criteria:**
  - **BIB and cosmic rays:** *looser* or $|t_{\text{jet}}| > 25$ ns;
  - **calorimeter noise:** *looser* or ($f_{hec} > 0.5$ and $|f_Q^{hec}| > 0.5$) or ($f_{em} > 0.95$ and $f_Q^{LAr} > 0.8$ and $|\eta| < 2.8$).
- ***medium* criteria:**
  - **BIB and cosmic rays:** *loose* or $|t_{\text{jet}}| > 10$ ns or ($f_{em} < 0.05$ and $f_{ch} < 0.1$ and $|\eta| < 2.0$) or ($f_{em} < 0.95$ and $f_{ch} < 0.05$ and $|\eta| < 2.0$);
  - **calorimeter noise:** *loose* or ($f_{hec} > 1 - |f_Q^{hec}|$) or ($f_{em} > 0.9$ and $f_Q^{LAr} > 0.8$ and $|\eta| < 2.8$).
- ***tight* criteria:**
  - **BIB and cosmic rays:** *medium* or ($f_{em} < 0.1$ and $f_{ch} < 0.2$ and $|\eta| < 2.5$) or ($f_{em} > 0.9$ and $f_{ch} < 0.1$ and $|\eta| < 2.5$) or ($f_{ch} < 0.01$ and $|\eta| < 2.5$) or ($f_{em} < 0.1$ and $|\eta| \geq 2.5$);
  - **calorimeter noise:** *medium* or ($f_Q^{LAr} > 0.95$) or ($f_{em} > 0.98$ and $f_Q^{LAr} > 0.05$)

It is important to note that in contrast with the electron or photon case, a jet satisfying the aforementioned criteria is classified as having "bad" quality.

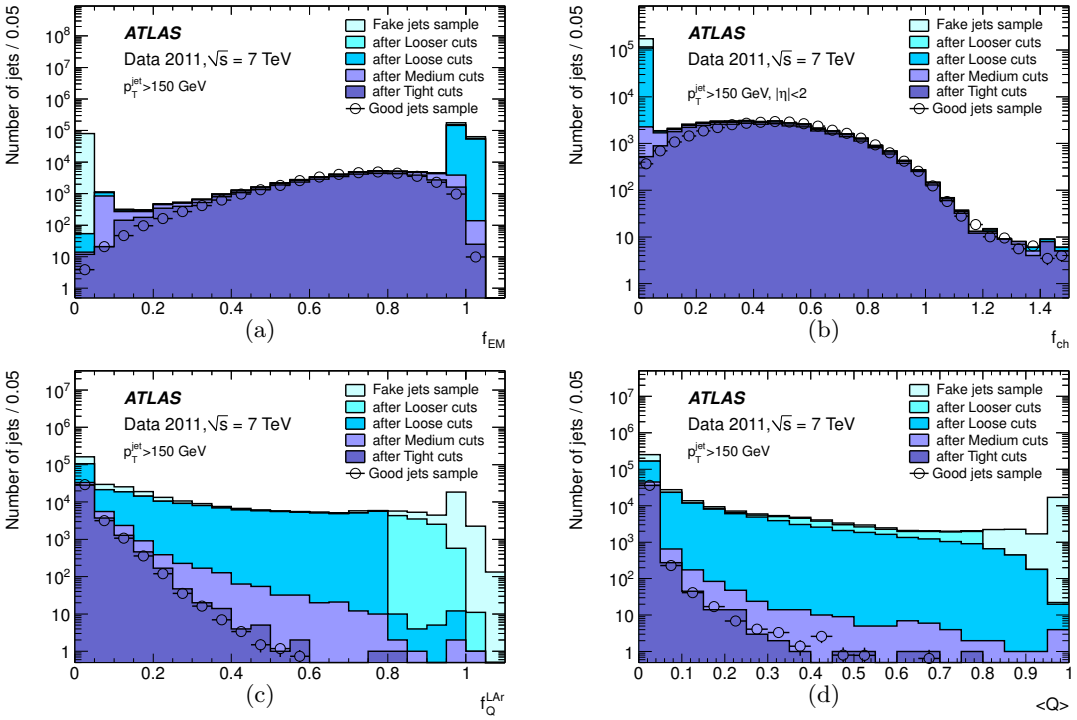


Figure 4.11: *Distributions of discriminating variables for a sample enriched in fake jets reconstructed using the 2011 $\sqrt{7}$ TeV dataset before and after applying the jet selection criteria. Distributions for a sample enriched in collision jets are also superimposed [76].*

4.5.3 Jet calibration

A different calibration scheme was used for 2011 and 2012 jets. For 2011 jets, the scheme, known as EM+JES, starts with jets reconstructed with EM topo-clusters and follows the steps shown in figure 4.12, while for 2012 jets, the scheme, known as LCW+GSC, starts with jets reconstructed with LCW topo-clusters and follows the steps shown in figure 4.13.

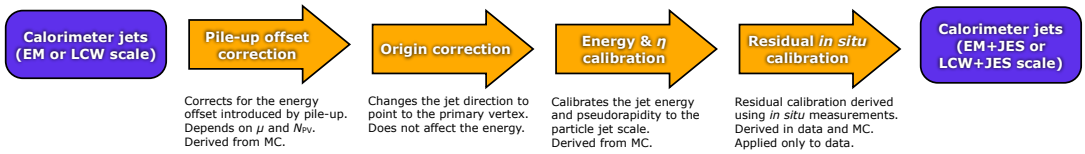


Figure 4.12: *Overview of the ATLAS jet calibration scheme used for the 2011 $\sqrt{s} = 7$ TeV dataset [74].*

In the following, additional details for each of the steps are given.

Origin correction

During reconstruction, topo-clusters are assigned a direction to complete their 4-vector in addition to its energy. The default choice is to point them at the centre of the detector. However, after reconstruction of the full event, a better assumption is that they originated from the position of the *first primary vertex*, defined as the vertex with the highest $\sum P_T^2$ of tracks. The origin correction accounts for this difference by finding the energy centre of the jet and modifying the jet 4-vector such that the energy is unchanged but the jet now originates from the *first primary vertex*. The η and ϕ resolution of jets before and after the origin correction is shown in figure 4.14. There is little change in the resolution in ϕ due to the small spread of the beam-spot in (x, y).

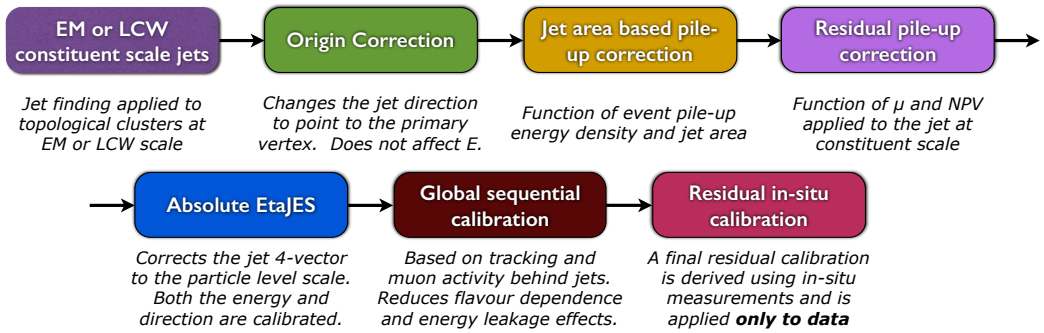
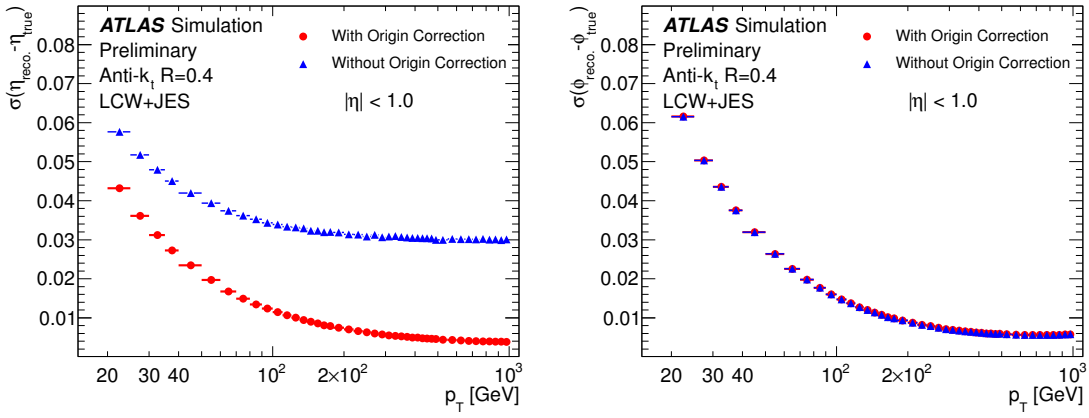


Figure 4.13: *Overview of the ATLAS jet calibration scheme used for the 2012 $\sqrt{s} = 8$ TeV dataset [77].*



(a) (b)

Figure 4.14: *The effect of the origin correction on the η resolution (a) and ϕ resolution (b) of $R = 0.4$ jets [77].*

Pile-up correction

The approach for 2011 jets consists of calculating the amount of transverse momentum generated by pile-up in a jet in MC simulations, and subtracting this offset from the reconstructed jet P_T. To characterise the in-time pile-up activity, the number of reconstructed primary vertices N_{pv} is used. For the out-of-time pile-up activity, the average number of interactions per bunch crossing ($\langle\mu\rangle$) at the time of the recorded event provides a good estimator. It is calculated using:

$$\langle\mu\rangle = \frac{L \times \sigma_{inel}}{N_{bunch} \times f_{LHC}}, \tag{4.9}$$

where L is the instantaneous luminosity per bunch, σ_{inel} is the total inelastic proton-proton cross section, N_{bunch} is the number of colliding bunches in the LHC and f_{LHC} is the LHC revolution frequency [78].

The MC-based jet calibration is derived by fitting the dependence on $N_{pv}(\langle\mu\rangle)$ for fixed values of $\langle\mu\rangle(N_{pv})$, and then averaging for the different fixed $\langle\mu\rangle(N_{pv})$. The P_T is corrected using

$$P_T^{corr} = P_T^{jet} - \alpha \times (N_{pv} - 1) - \beta \times \langle\mu\rangle, \tag{4.10}$$

where α and β are the parameters obtained with the MC-based method previously described.

Both parameters are jet size, jet algorithm and η dependent.

For 2012 jets, an area based subtraction method was employed [79, 80]. The area of a jet is calculated using FastJet using an active area algorithm in which ghost particles (particles with infinitesimal P_T) are uniformly added to an event before the event is reclustered. The number of ghosts clustered into each jet then gives a measure of the area of the jet as seen in figure 4.15a. The pile-up energy density of the event is calculated using jets reconstructed with the k_T algorithm and $R = 0.4$ in the central ($|\eta| < 2.0$) region. The energy density of each jet is defined as P_T/A and the event P_T density (ρ) is defined as the median energy density of these jets and it is shown in figure 4.15b for events with different N_{pv}.

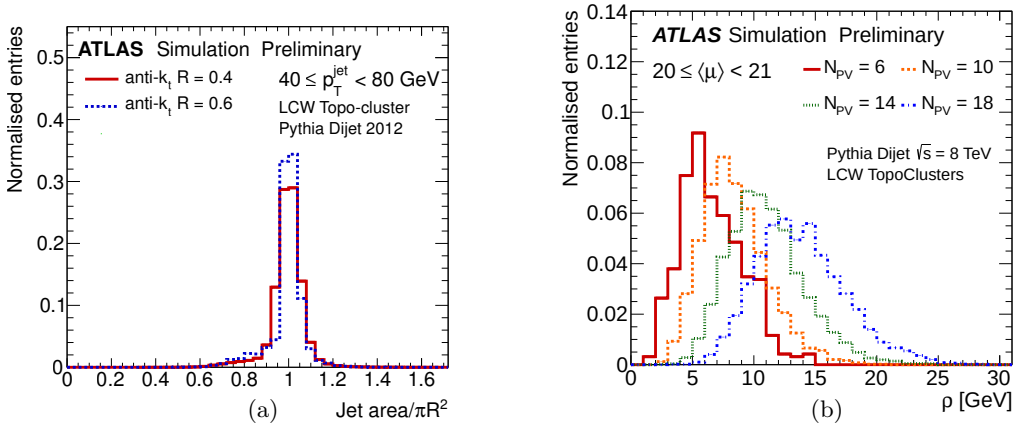


Figure 4.15: *Jet active area for the anti-k_T algorithm with different values of R (a) and ρ for an average number of interactions $20 < \langle\mu\rangle < 21$ for four different values of N_{pv} [80].*

It is observed that after this correction there remains some small dependence of the jet P_T on pile-up so an additional residual correction is applied using the same procedure as for the 2011 jets, that is, parametrising the residual correction in terms of N_{pv} and $\langle\mu\rangle$ and obtaining the required parameters from MC simulations.

4.5.4 Jet energy scale (JES)

The jet energy scale calibration is derived as a correction which relates the reconstructed jet energy to the truth jet energy. The JES factors are derived from isolated jets from an inclusive jet Monte Carlo sample.

After the calibration in energy it is found that in particular regions of the detector there is a bias in the η distribution with respect to the truth jets. Therefore, an additional correction in purely the angle of the jet is applied to resolve this bias.

4.5.5 Global sequential correction (GSC)

At this point of the calibration, it is observed that there is a difference between the closure of quark and gluon initiated jets (as defined by angular matching to partons in Monte Carlo events) whereby a response difference of up to 8% is observed between quark and gluon initiated jets [74].

This difference resulted in one of the largest uncertainties on the jet energy calibration for 2011 jets. To reduce the difference between the jet responses to quark and gluons and thereby improve both jet resolution and jet energy scale uncertainties, further corrections are applied for 2012 jets following the JES calibrations. These corrections include a "punch-through" correction to correct high P_T jets whose energy is not fully contained within the calorimeter jet.

The corrections are applied depending on the topology of energy deposits in the calorimeter, tracking information and muon spectrometer information. There are three corrections applied to LCW calibrated jets that are based on:

- the number of tracks with $P_T > 1$ GeV associated to the jet;
- the P_T-weighted transverse width of the jet measured using tracks with $P_T > 1$ GeV associated to the jet;
- the amount of activity behind the jet as measured in the muon spectrometer.

The corrections are derived from MC simulations and are η dependent [81].

In-situ jet energy calibration

After the MC-based calibration of jets, in-situ techniques employing the balance of physics objects in the transverse plane, are used in the final stage of the calibration. The P_T of reference objects (photons, Z bosons or other jets) and the jets being calibrated are compared in both data and MC to measure the ratio

$$\frac{\mathcal{R}_{data}}{\mathcal{R}_{MC}} = \frac{\langle P_T^{jet}/P_T^{ref}\rangle_{data}}{\langle P_T^{jet}/P_T^{ref}\rangle_{MC}}, \tag{4.11}$$

This quantity defines a residual correction which is applied to jets reconstructed in data.

There are three steps in the in-situ corrections:

- firstly, dijet events are employed to apply an η-intercalibration in which the average P_T for forward jets ($0.8 \leq |\eta| \leq 4.5$) is equalised to the p_{T} of balancing jets in the central region ($|\eta| < 0.8$) [82] and aims to remove any residual η dependence in the jet response following the MC calibration. The η-intercalibration correction factors are generally below 2%;
- after the η-intercalibration, the balance of Z bosons or photons recoiling against jets is used to derive in-situ JES corrections for jets with $|\eta| < 0.8$;
- finally, high P_T jets are calibrated using events in which a system of low-p_{T} jets recoils against a single high-p_{T} jet (multi-jet balance). This requires the low-p_{T} system to be well calibrated and so the method can be iterated starting from jets already well calibrated by other in-situ methods and increasing in P_T until the limited number of events prevents accurate calibration.

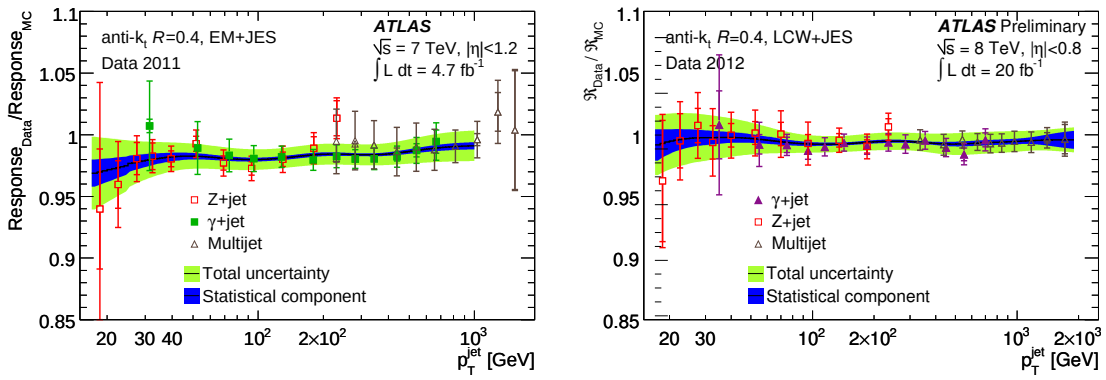
Figure 4.16 shows $\mathcal{R}_{data}/\mathcal{R}_{MC}$ for 2011 and 2012 jets.

4.5.6 B-tagging

The ability to identify jets containing b-hadrons was needed for the high transverse-momentum top quark analysis to take advantage of the fact that all the $t\bar{t}$ decay modes contain b-quarks in the final state.

All the b-tagging algorithms are based on MC predictions for the signal (b-jet) or background (light- or in some cases c-jet) hypotheses.

The algorithm used in this dissertation is called MV1, which is a neural-network based b-tagging algorithm that combines the predictions from three other algorithms, IP3D, SV1 and JetFitterCombNN [83]. They are briefly described below.

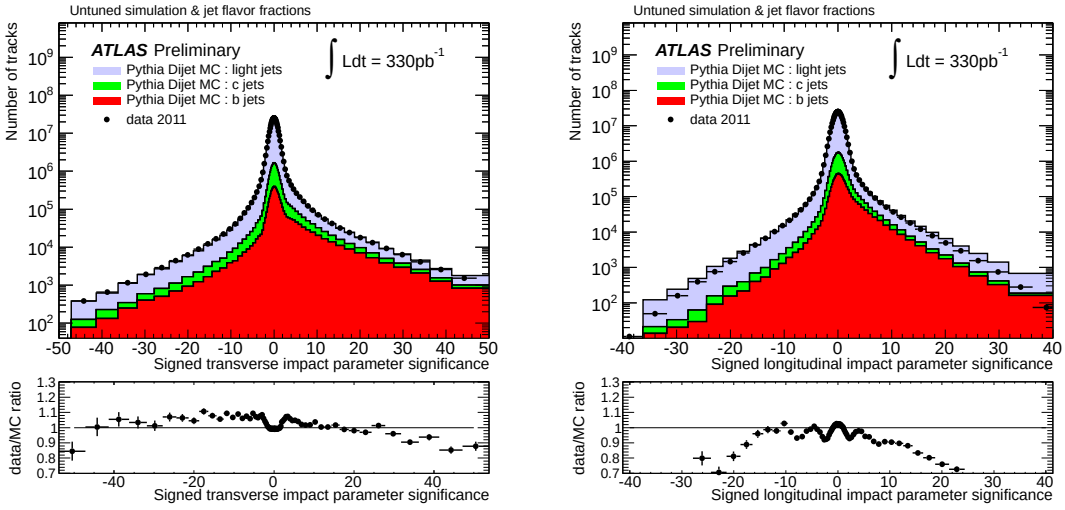


(a) (b)

Figure 4.16: *Ratio of response measured in data to response measured in MC for Z+jet, γ+jet and multi-jet balance in-situ analyses for 2011 (a) [74] and 2012 (b) [77] jets. The black line is the combined correction with its associated uncertainty.*

IP3D

The IP3D high-performance tagging algorithm uses the impact parameter significances of all the tracks in a jet as the basis of the method. The transverse impact parameter d_0 is the distance of closest approach of the track to the primary vertex while the z coordinate of the track at that point is referred to as z_0 and often called longitudinal impact parameter. On the basis that the decay point of the b-hadron must lie along its flight path, the impact parameter is signed to further discriminate the tracks from b-hadron decays from tracks originating from the primary vertex. The sign is positive if the track extrapolation crosses the jet direction in front of the primary vertex and negative otherwise. Therefore, tracks from b/c hadrons tend to have a positive sign. The significances of both variables d_0/σ_{d0} and z_0/σ_{z0} , which gives more weight to tracks measure precisely, are the main ingredients of IP3D, and are shown in figure 4.17.



(a) (b)

Figure 4.17: *Distribution of the signed significance of the transverse (a) and longitudinal (b) impact parameters with respect to the primary vertex for tracks of b-tagging quality associated to jets [84].*

IP3D uses a likelihood ratio technique in which the input variables are compared to pre-defined smoothed and normalised distributions for both the b- and light jet hypotheses, obtained from MC simulation. The distributions in this case are two dimensional histograms of d_0/σ_{d0} and z_0/σ_{z0}.

SV1

The SV1 secondary vertex algorithm uses the inclusive vertex formed by the decay products of the b-hadron, including the products of the eventual subsequent charm hadron decay. The search starts by building all two-track pairs that form a good vertex, using only tracks associated to the jet and far enough from the primary vertex. Vertices compatible with a V^0 or material interaction are rejected. All tracks from the remaining two-track vertices are combined into a single inclusive vertex, using an iterative procedure to remove the worst track until the χ^2 of the vertex fit is considered to be good enough.

The discriminating variables used are the invariant mass of all tracks associated to the vertex, the ratio of the sum of the energies of the tracks in the vertex to the sum of the energies of all tracks in the jet, and the number of two-track vertices. These variables are combined using a likelihood ratio technique.

JetFitterCombNN

The JetFitterCombNN algorithm is a neural network based combination of the IP3D and the JetFitter algorithm.

The JetFitter exploits the topology of weak b- and c-hadron decays inside a jet. A Kalman filter is used to find a common line on which the primary vertex and the b- and c-vertices lie, as well as their position on this line, giving an approximated flight path for the b-hadron. The discrimination between b-, c- and light jets is based on a likelihood using similar variables as in the SV1 tagging algorithm and additional variables such as the flight length significances of the vertices [85].

Chapter 5

Identification of high transverse-momentum top quarks reconstructed as single jets in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector

In this chapter a new method to identify highly boosted massive particles based on their hadronic decay modes is presented. The method is based on the reconstruction of the so-called large-R jets and uses substructure techniques to identify the massive particles, reduce the background and study their properties. The techniques were tested via the identification of boosted top quarks.

5.1 Event selection

The initial event selection is referred to as pre-selection and is based on the $t\bar{t}$ decay properties.

5.1.1 Pre-selection

This analysis makes use of the semi-leptonic decay mode (see section 1.4.2), in which one W boson decays into an electron and a neutrino and the other decays hadronically. The semi-leptonic channel is the preferred channel for this analysis as it provides a clear signature while at the same time keeping the possibility to recover all the decay products of the top quark that decayed hadronically and, therefore, to reconstruct the top quark directly.

At high transverse momentum, the massive, hadronically decaying top quark will be produced with a significant boost and will tend in turn to produce collimated decays. In such a case, the hadronic products are highly collimated and resolving such hadrons into different jets is difficult. Therefore it is more convenient to reconstruct these hadrons as a single jet using a large value for the distance parameter in the anti-k_{T} jet algorithm, the so-called "large-R" jet, and use substructure techniques to study its properties.

Using this information, the pre-selection was developed based in the presence of one high energetic isolated electron and the presence of a high p_{T} large-R jet. In an effort to fully establish the boosted top quark identification capabilities of the substructure-analysis method, no addi-

tional requirements, such as the presence of missing energy or b-tagging related requirements were included in the pre-selection. The pre-selection criteria were:

- three electron trigger chains (see section 4.1) were used to select suitable candidates for semi-leptonic top decays in the electron channel. Different chains were used for different periods of data taking to take advantage of the data taking conditions:
  - EF_e20_medium was required for periods A to J. This chain consists of a high-level electron trigger with a transverse energy threshold of 20 GeV seeded by the L1_EM14 trigger, with a 14 GeV threshold. It involves medium high-level trigger electron identification criteria;
  - EF_e22_medium was required for period K. This chain consists of a high-level electron trigger with a transverse energy threshold of 22 GeV seeded by the L1_EM16 trigger, with a 16 GeV threshold. It involves medium high-level trigger electron identification criteria;
  - EF_e22vh_medium1 or EF_e45_medium1 for periods I to M. The EF_e22vh_medium1 uses the same thresholds than the EF_e22_medium chain, but with an isolation requirement using the 4 hadronic towers behind the core cluster at the L1 level. The EF_e45_medium1 chain consists of a high-level electron trigger with a transverse energy threshold of 45 GeV seeded by the L1_EM30 trigger, with a 30 GeV threshold. Both chains involve the medium high-level electron trigger identification criteria.
- a primary vertex, with at least 5 tracks associated to it with a minimum p_{T} of 400 MeV, was required to suppress non-collision background;
- electrons with $E_T > 40$ GeV and $|\eta| < 2.47$ were selected (see section 4.3.1):
  - isolation was required both within the electromagnetic calorimeter and the inner detector (see section 4.3.3). Both cuts were chosen as to maintain a reconstruction efficiency of 90%;
  - quality was ensured by imposing the *tight* identification criteria (see section 4.3.2);
  - candidates in the crack region of the calorimeter, $1.37 < |\eta| < 1.52$, were rejected;
  - candidates were required to match a trigger object at the EF level; the two objects were considered matched if the distance between them in the $\eta - \phi$ plane was less than 0.2;
  - candidates in data samples were calibrated using the procedure described in section 4.3.4. A smearing in energy was applied to electron candidates in MC simulated events instead. An extra correction for candidates with $|\eta|$ between 1.52 and 1.57 was applied, which was added to the procedure after an underestimation of the energy for electrons reconstructed in this region was found;
- jets were reconstructed using the anti-k_{T} algorithm with $R = 0.4$ (see section 4.5.1) and calibrated using the EM+JES scheme (see section 4.5.3). Events with at least one jet passing the *looser* identification criteria (see section 4.5.2) were rejected. The overlap with electrons was taken care of by removing the jet closest to a reconstructed electron if their distance in the $\eta - \phi$ plane was smaller than 0.2. For this requirement, electrons with $E_T > 25$ GeV were used;
- events were required to have exactly one final electron. Electrons close to a jet (distance in the $\eta - \phi$ plane between them smaller than 0.4) were not considered. For this step, the jets surviving the previous conditions were used;
- large-R jets were reconstructed using the anti-k_{T} algorithm with $R = 1.5$ and LCW topo-clusters. Events were required to have at least one large-R jet with $p_{\mathrm{T}} > 400$ GeV. The choice of the value of the R parameter is justified in section 5.2.

5.1.2 Data selection

The pre-selection was applied to the data collected during the 2011 LHC running period at a centre-of-mass energy of 7 TeV (see section 2.2). Only good quality data in which every part of the ATLAS detector was fully operational were used. The resulting dataset corresponds to an integrated luminosity of approximately 4.6 fb^{-1}. There are 2817 events in the pre-selected sample.

Figures 5.1 and 5.2 show two events that pass the pre-selection requirements. In both events the presence of an electron (green line), and of a large amount of energy deposited in the calorimeters (yellow squares) that will be reconstructed into at least one large-R jet with $p_{\mathrm{T}} > 400$ GeV, can be seen.

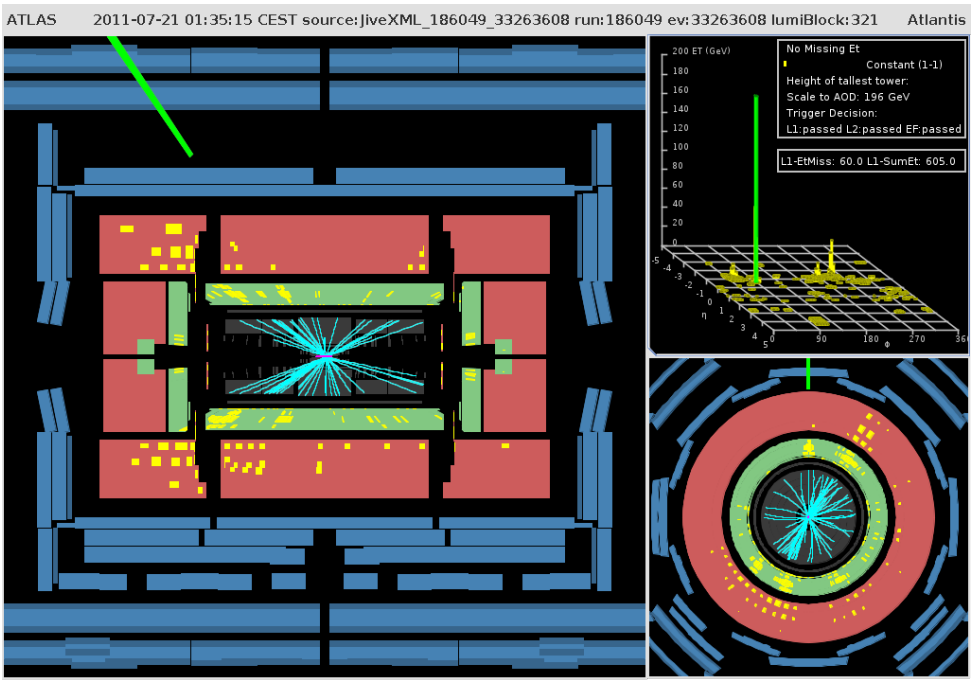


Figure 5.1: *Event display showing an event from the 2011 $\sqrt{s} = 7$ TeV dataset surviving the pre-selection requirements. Green lines represent isolated electrons of tight quality, light blue lines represent tracks in the inner detector and yellow squares represent energy deposits in the calorimeters. The EM calorimeter is drawn in green while the hadronic calorimeter is drawn in red.*

Figure 5.3 shows the data distributions after pre-selection as functions of the transverse momentum of the leading large-R jet ($p_{\mathrm{T}}^{\mathrm{jet1}}$), the transverse energy of the electron (E_T^e), the rapidity of the large-R jet ($|y^{\mathrm{jet1}}|$), the pseudo-rapidity of the electron ($|\eta^e|$) and the azimuthal angle of the leading large-R jet (ϕ^{jet1}) and that of the electron (ϕ^e).

5.1.3 MC selection

The pre-selection was applied to the signal and background samples detailed in section 3.2.1, with the exception of the QCD dijet samples. The background processes which are able to mimic the properties of a semi-leptonic $t\bar{t}$ decay are:

- **W+jets production:** the $W \to e\nu$ process includes an isolated electron while extra hard partons can provide the large-R jet;
- **Z+jets production:** the $Z \to ee$ process may produce a single isolated electron signature if one of the electrons is reconstructed and selected and the other one is not. As with the

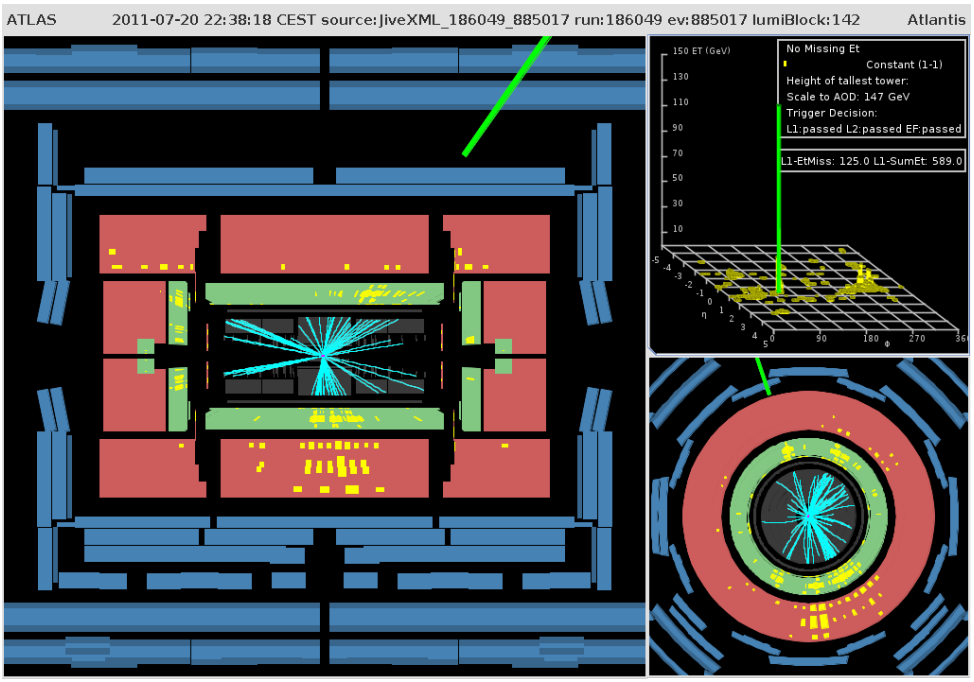


Figure 5.2: *Event display showing an event from the 2011 $\sqrt{s} = 7$ TeV dataset surviving the pre-selection requirements. Green lines represent isolated electrons of tight quality, light blue lines represent tracks in the inner detector and yellow squares represent energy deposits in the calorimeters. The EM calorimeter is drawn in green while the hadronic calorimeter is drawn in red.*

W boson background, extra hard partons can effectively fake a semi-leptonic $t\bar{t}$ event;

- **Single top:** the production of single top in any of its modes has all the necessary ingredients to fake semi-leptonic $t\bar{t}$ events. In the s- and t-channels, the top quark may decay leptonically, providing an isolated electron while in the Wt-channel the single top or the additional W may produce the isolated electron signature.

- **QCD multi jet production:** Although it is very difficult for a jet to mimic an isolated electron, the huge cross section of multi-jet events through QCD interactions could fake the signal.

In addition to the pre-selection, the following corrections were applied:

- although the MC samples used for this analysis contained all the different data taking conditions that were present during 2011 data taking, it was necessary to re-weight the MC samples so that the fraction of those different conditions was the same in both data and MC. The re-weighting was done both to account for the pileup conditions, based on $\langle\mu\rangle$, and for the presence and evolution of a calorimeter coverage issue that was present during several periods of data-taking;

- scale factors were applied to the events to compensate for residual differences in efficiency between data and MC, both for the electron identification and the trigger.

QCD background estimation

Due to the nature of the requirements for the pre-selection, mainly the presence of one isolated electron, it is not possible to obtain enough statistics with the available MC dijet

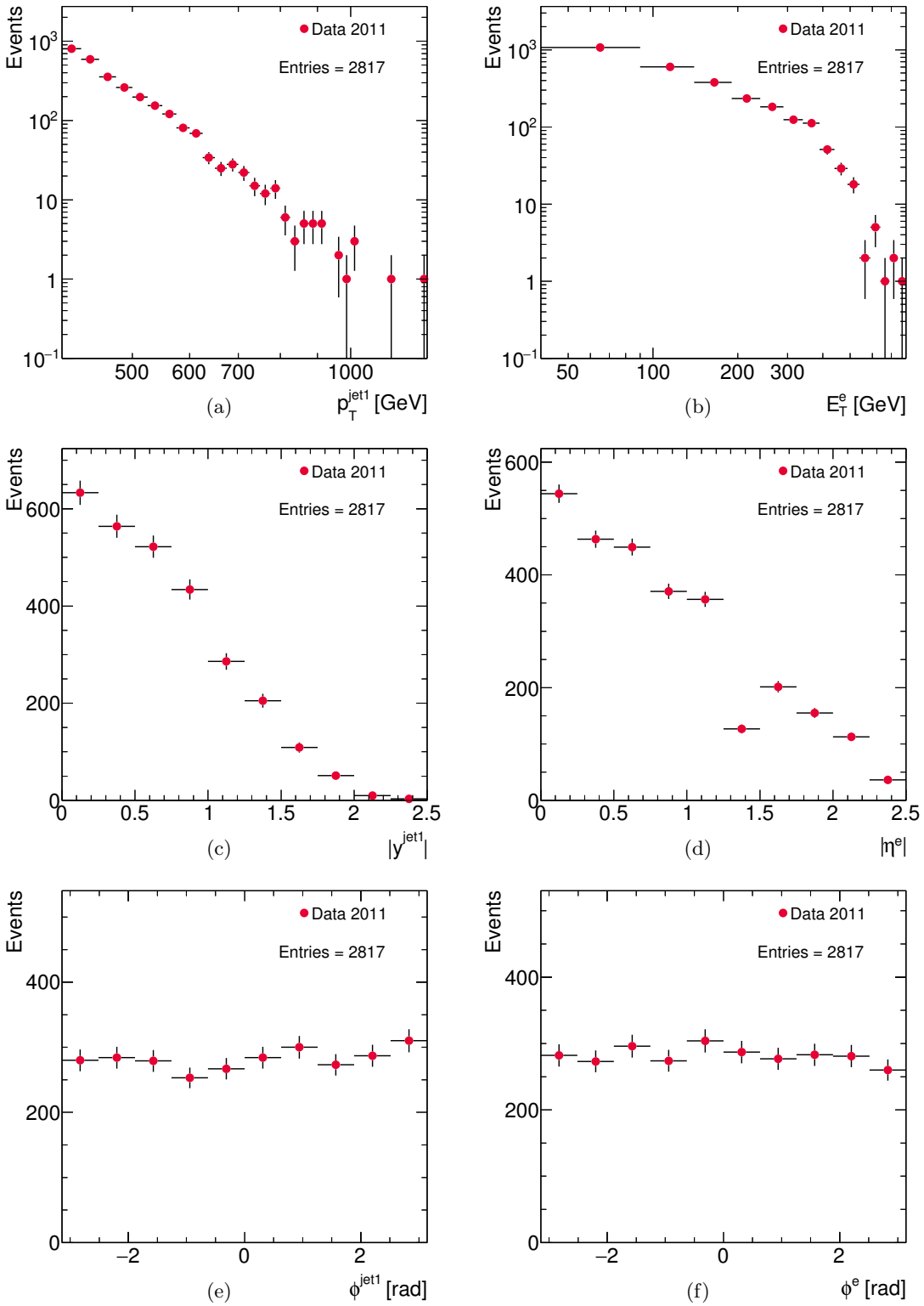


Figure 5.3: *Data distributions after pre-selection as functions of* $p_{\mathrm{T}}^{\mathrm{jet1}}$ *(a),* E_{T}^{e} *(b),* $|y^{\mathrm{jet1}}|$ *(c) ,* $|\eta^{e}|$ *(d),* ϕ^{jet1} *(e) and* ϕ^{e} *(f).*

samples. To improve the statistics of the MC dijet sample, the jet-electron model was used instead. This method estimates the QCD background by selecting events with one jet that mimics most of the properties of the electron. Although this method was originally designed for its implementation on data events, the high pre-scale of the trigger chains used makes it impossible to use it on the 2011 data and thus it was applied to MC dijet events instead to obtain the QCD background. The pre-selection used differs from the one defined in section 5.1.1 as explained in the following:

- events were required to pass any of the following jet and forward jet trigger chains (see section 4.2):
  - EF_j20_a4_EFFS, EF_j20_a4tc_EFFS, EF_j10_a4tc_EFFS, EF_j15_a4tc_EFFS, EF_fj20_a4tc_EFFS, EF_fj10_a4tc_EFFS, EF_fj15_a4tc_EFFS;
- events with exactly one jet-electron with $p_{\mathrm{T}} > 40$ GeV and $|\eta| < 2.47$ were selected. Jet-electrons are a construct that represents jets that are nearly miss-reconstructed as electrons and were defined as follows:
  - jets reconstructed using the anti-k_{T} algorithm with $R = 0.4$ and calibrated using the EM+JES scheme were used as the starting point;
  - the p_{T} at the electromagnetic scale was required to be higher than 40 GeV;
  - the jet-electron candidates in the crack region were rejected ($1.37 < |\eta| < 1.52$);
  - the electromagnetic fraction was required to be higher than 0.8 and lower than 0.95;
  - the number of tracks with $p_{\mathrm{T}} > 1$ GeV associated to the jet-electron candidate was required to be at least 4, to reduce the amount of converted photons;
- events with at least one selected electron were vetoed. Electrons were selected following the same rules defined for the nominal pre-selection with the exception of the isolation, which was not required, and the quality requirement, which were loosened from *tight* to *loose* (see section 4.3.2).

MC distributions after pre-selection

Figure 5.4 shows the MC distributions after pre-selection for the signal and the most important backgrounds as functions of $p_{\mathrm{T}}^{\mathrm{jet1}}$, E_T^e, $|\eta^{\mathrm{jet1}}|$, $|\eta^e|$, ϕ^{jet1} and ϕ^e. The distributions have been normalised to unit area.

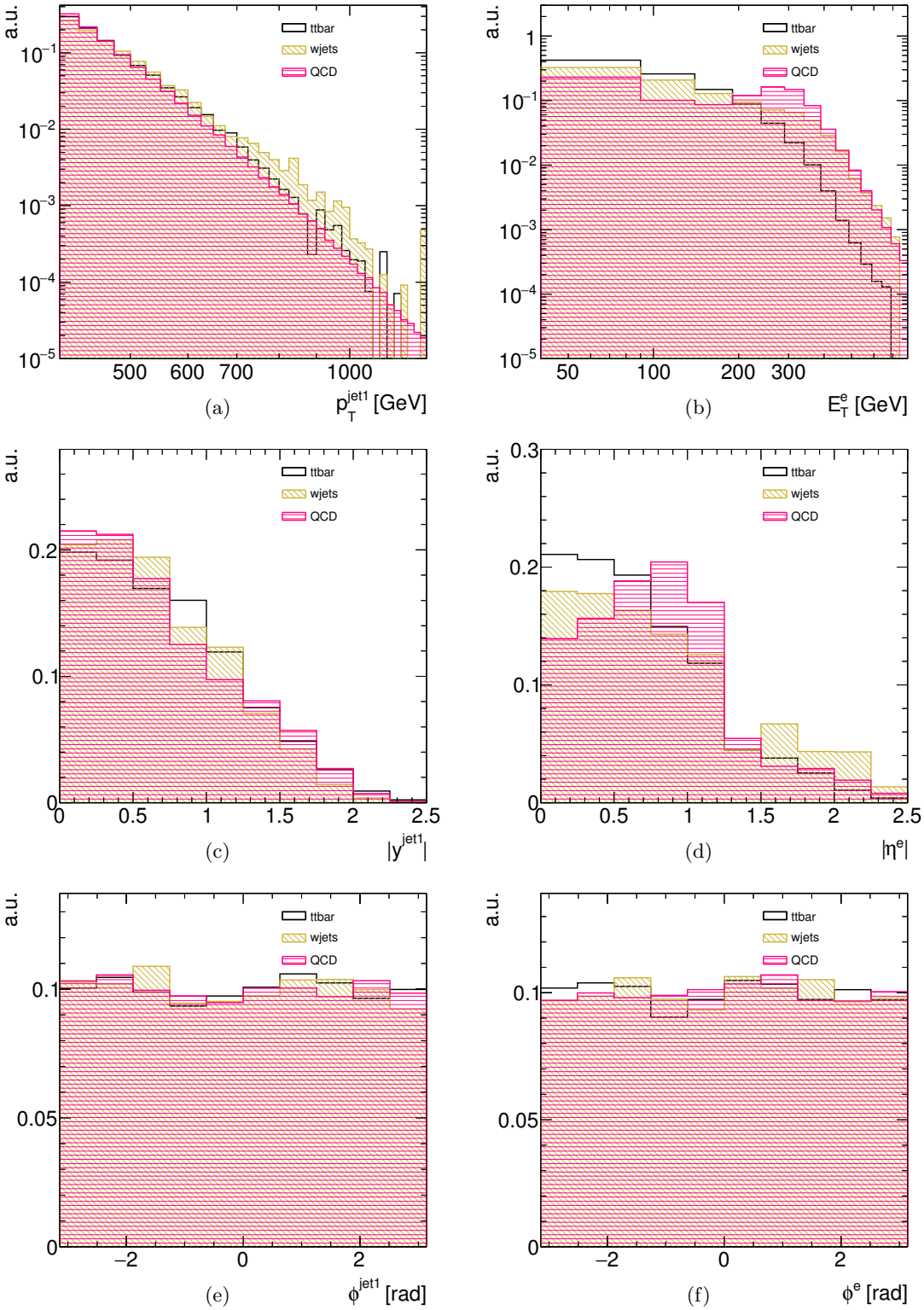


Figure 5.4: *MC distributions after pre-selection for the signal and the most important backgrounds as functions of p_T^{jet1} (a), E_T^e (b), $|y^{jet1}|$ (c), $|\eta^e|$ (d), ϕ^{jet1} (e), and ϕ^e (f). The distributions are normalised to unit area*

5.2 Substructure analysis method

Once candidates for $t\bar{t}$ semi-leptonic production have been obtained using the pre-selection, it is expected that in a $t\bar{t}$ event, the hadronically decayed top products will be contained within one of the required large-R jets. The leading large-R jet was selected for the substructure analysis if its distance in the $\eta - \phi$ plane to the selected electron was larger than 1.5. Otherwise, the sub-leading large-R jet was selected instead and was required to have $p_{\mathrm{T}} > 400$ GeV. This choice is based on the assumption that the electron coming from the leptonically decaying top quark is well separated from the collimated products of the hadronically decaying top quark.

The main idea behind the method proposed in this analysis for studying the substructure of large-R jets is to combine the characteristics of both the k_{T} (well defined subjets and sensitivity to soft radiation) and the anti-k_{T} (better identification of the hard component of the event) algorithm.

Subjet reconstruction was performed by applying the k_{T} algorithm on the constituents of the large-R jet to identify structures within the large-R jet which correspond to the decay products of the hadronically decaying top quark. The kinematic properties of the boosted particle reconstructed using the subjets in the large-R jet are affected by contributions from soft QCD radiation, and, due to the running conditions in the LHC, pile-up contributions. To remove the influence of these effects in an infrared-and-collinear safe way and recover the kinematic properties of the decay products, the anti-k_{T} algorithm was used to characterise the decay products. The soft-resilient properties of the anti-k_{T} (see section 1.3.3) make the removal of the soft contribution a natural outcome of this method.

The following procedure, detailed schematically in figure 5.5, was applied to every event satisfying the pre-selection cuts, both for the data and the MC samples, yielding exactly one reconstructed top quark candidate in each event:

- the constituents of the large-R jet chosen during the pre-selection were re-clustered using the k_{T} algorithm. For this step, the algorithm was ran in exclusive mode with 3 jets as the target (see section 1.3). This procedure always yields three subjets from the parent large-R jet (see figure 5.5a);
- the constituents of the three subjets were re-clustered using the anti-k_{T} algorithm with $R = 0.4$, once for each list of constituents (see figure 5.5b). The leading jet of each clustering process was taken as the final object, referred to as "small-R jet" (see figure 5.5c);
- the reconstructed top was defined as the four-vectorial sum of the three final small-R jets.

It is important to note that there is nothing specific to the top quark decay properties included in the method, apart from the fact that it is expected to have a 3-subjet structure. The method can be applied to any massive particle decaying into a single large-R jet simply by changing the target number of jets in the first step of the procedure to the expected number of subjets.

Figure 5.6 shows the invariant mass of the three small-R jets (M_{JJJ}), the invariant mass of each of the three possible pairs of small-R jets (M_{JJ}), the fraction between the p_{T} of the leading small-R jet and the p_{T} of the parent large-R jet ($p_{\mathrm{T}}^{\mathrm{J1}}/p_{\mathrm{T}}^{\mathrm{jet1}}$) and finally $\sqrt{\mathrm{dmin}_{21}}$ that represents the scale at which the k_{T} algorithm would merge two subjets into one for the MC samples. The M_{JJJ} distribution for the $t\bar{t}$ MC sample peaks around M_{top} as expected, whereas the background MC samples show a very different behaviour. Similarly, the M_{JJ} distribution for the $t\bar{t}$ MC sample shows a peak, around M_W, as expected. Therefore, this method gives a reliable reconstruction of the mass of the underlying particles. The $p_{\mathrm{T}}^{\mathrm{J1}}/p_{\mathrm{T}}^{\mathrm{jet1}}$ distribution takes lower values for the $t\bar{t}$ MC sample than for the background MC samples while the $\sqrt{\mathrm{dmin}_{21}}$ distribution takes higher values for the $t\bar{t}$ MC sample than for the background MC samples.

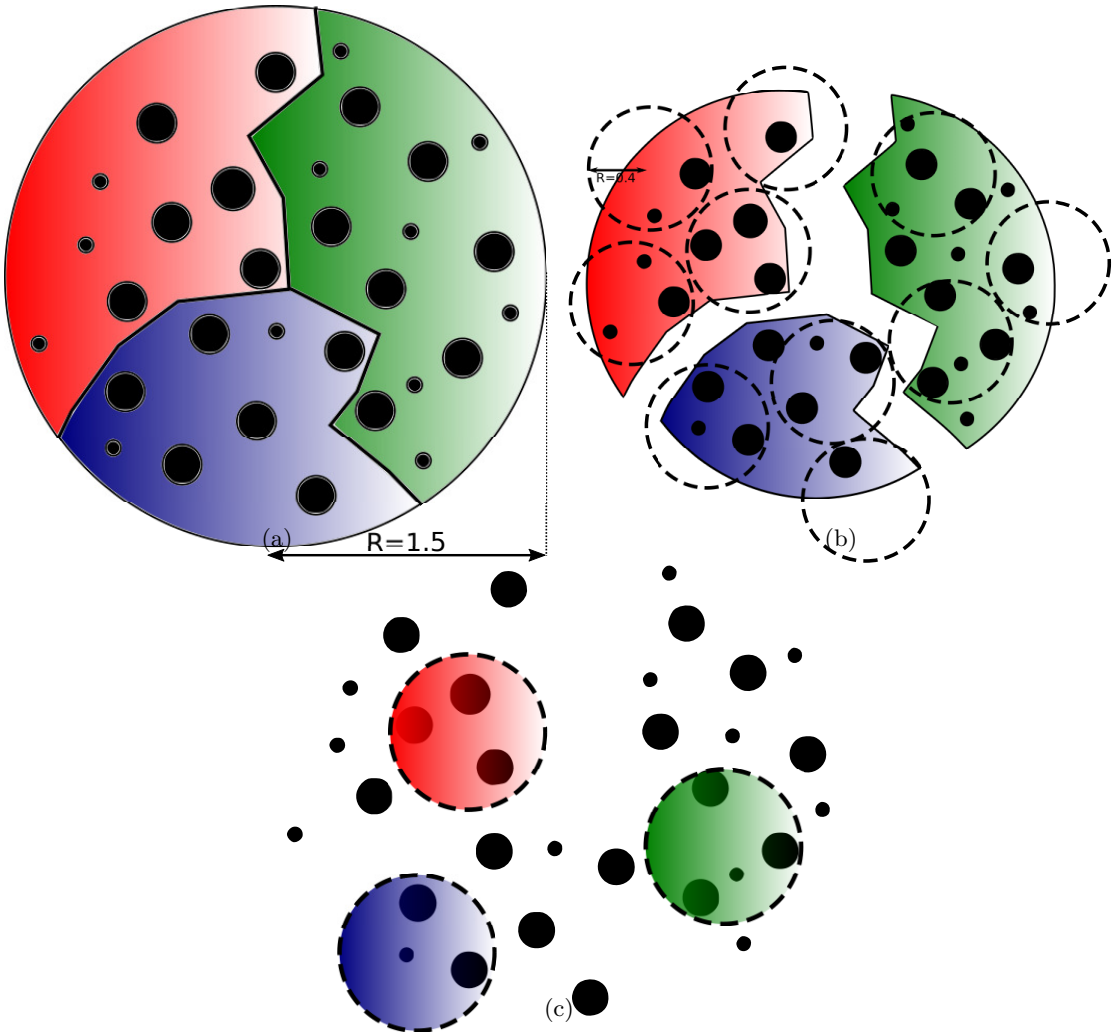


Figure 5.5: *Scheme of the substructure analysis method.*

Figure 5.7a shows the M_{JJJ} distributions obtained by using $R = 1$ and $R = 1.5$ for the distance parameter used in the large-R jet reconstruction for $t\bar{t}$ MC events. Clearly, a large number of events is lost in the $R = 1$ sample compared to the nominal $R = 1.5$ sample since many hadrons lie outside the cone and so the measured invariant mass is smaller than M_{top} for a significant fraction of events. This demonstrates that a larger R value is preferred to ensure that all the top decay products are contained within the large-R jet.

Figure 5.7b shows the M_{JJJ} distributions for $t\bar{t}$ events before and after applying the anti-k_{T} algorithm to the small-R jets: a narrow peak around the mass of the top is observed for the mass reconstructed from the anti-k_{T} jets. This demonstrates that this method effectively removes the unwanted soft contributions.

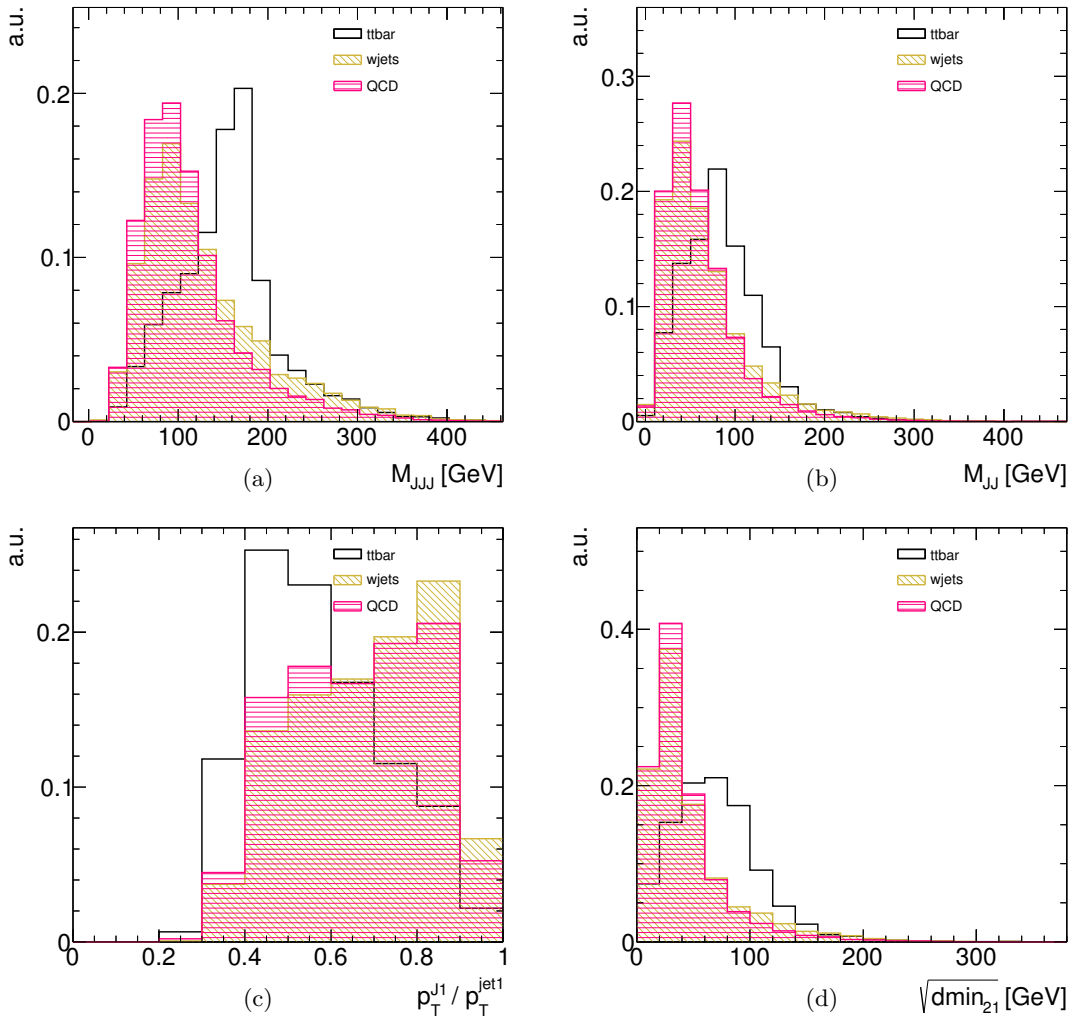


Figure 5.6: *MC distributions after pre-selection for the signal and the most important backgrounds as functions of* M_{JJJ} *(a),* M_{JJ} *(b),* $p_{\mathrm{T}}^{\mathrm{J1}}/p_{\mathrm{T}}^{\mathrm{jet1}}$ *(c) and* $\sqrt{\mathrm{dmin}_{21}}$ *(d). The distributions are normalised to unit area.*

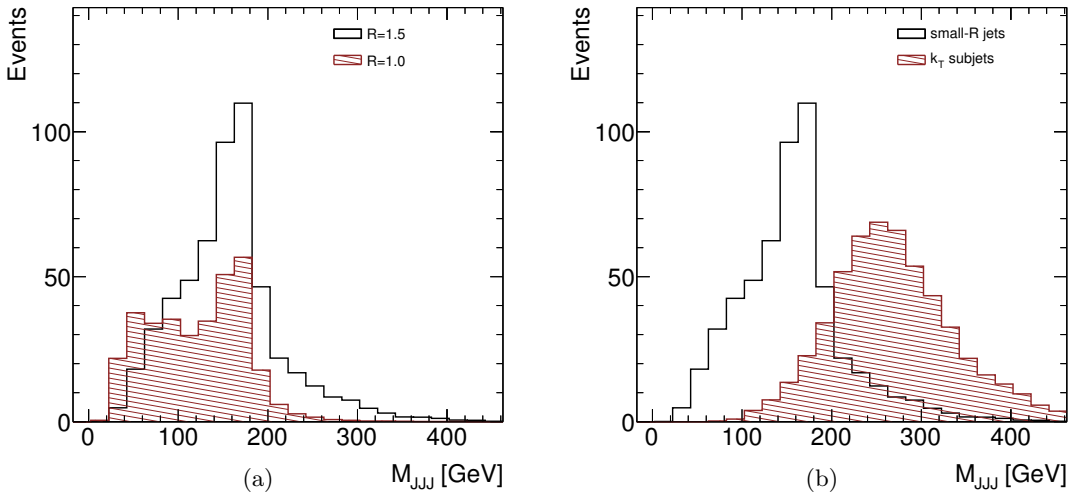


Figure 5.7: *MC distribution for the* $t\bar{t}$ *sample after pre-selection as a function of* M_{JJJ} *obtained using* $R = 1$ *and* $R = 1.5$ *as the distance parameter for the large-R jet reconstruction (a). MC distribution for the* $t\bar{t}$ *sample after pre-selection as a function of* M_{JJJ} *obtained using the small-R jets and the nominal* k_{T} *subjets (b).*

5.3 MC normalisation

Once the pre-selection was applied, the contribution of each MC sample, both from signal and background, was obtained by fitting the resulting histograms to the data to obtain the correct fraction of each type of process as dictated by the shape of the distributions in the data. The fit was made simultaneously on two variables, via a χ^2 minimisation method which takes into account the correlations between both variables. The variables chosen were the p_{T} of the sub-leading large-R jet ($p_{\mathrm{T}}^{\mathrm{jet2}}$) and the invariant mass of the sub-leading and sub-sub-leading small-R jets ($M_{\mathrm{J2,J3}}$). These variables were found to be sensitive to the differences between signal and background and between the different sources of background. The normalised distribution for those variables after pre-selection are shown in figure 5.8. In figure 5.9 the correlation between the variables used in the fit for both the $t\bar{t}$ MC samples and the $W+jets$ MC samples is shown. It is observed that the variables chosen are not strongly correlated.

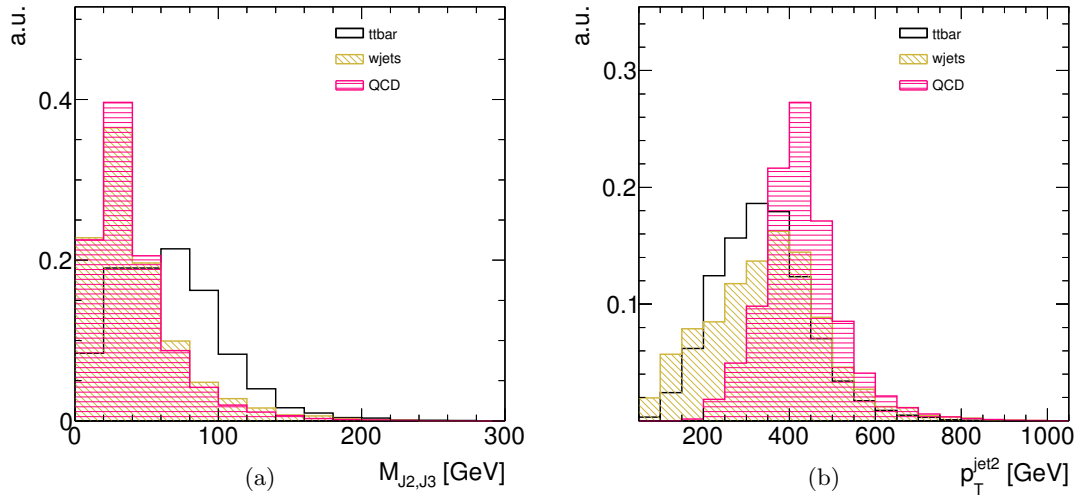


Figure 5.8: *MC distributions after pre-selection for the signal and the most important backgrounds as functions of $M_{\mathrm{J2,J3}}$ (a) and $p_{\mathrm{T}}^{\mathrm{jet2}}$. The distributions are normalised to unit area.*

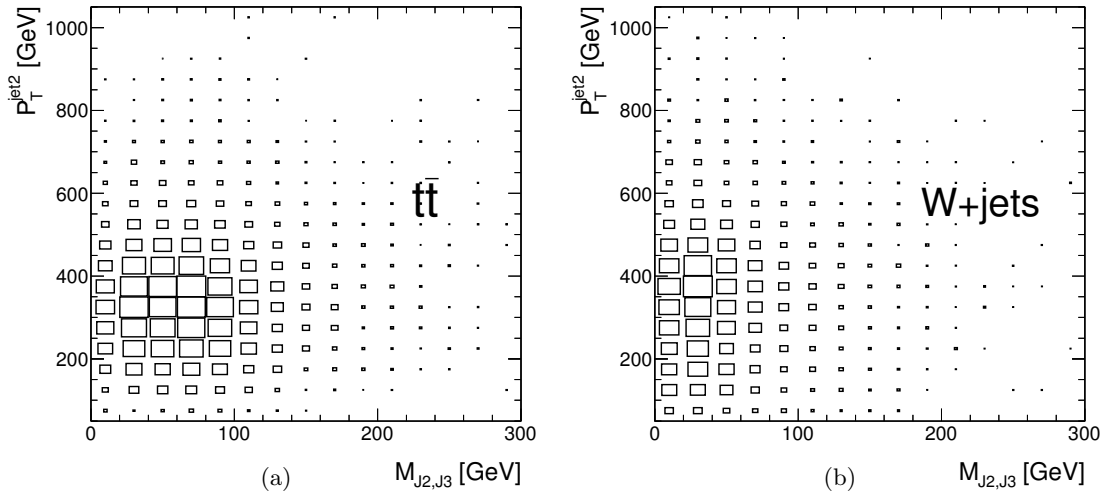


Figure 5.9: *Correlation between $M_{\mathrm{J2,J3}}$ and $p_{\mathrm{T}}^{\mathrm{jet2}}$ for $t\bar{t}$ MC samples (a) and W+jets MC samples (b).*

The overall MC distribution was then defined by summing three contributions:

- the contributions from W+jets and Z+jets (only one free parameter, α) are added accord-

ing to their predicted cross sections;

- the contributions from $t\bar{t}$ and single top (only one free parameter, β) are added according to their predicted cross sections;
- the contribution from QCD multijet events (γ).

The optimal values for these three parameters were then obtained from a χ^2 to extract the fractions of the different processes present in the final admixture. The values obtained and the corresponding fractions are shown in Table 5.1. The reduced χ^2 obtained from the fit was 0.64, which indicates that the fit is reasonably good.

Sample	Fit parameter	Percentage after pre-selection
$t\bar{t}$	0.9 ± 0.1	$18 \pm 2\%$
Single top	0.9 ± 0.1	$2.7 \pm 0.3\%$
W+jets	0.94 ± 0.06	$68 \pm 1\%$
Z+jets	0.94 ± 0.06	$9.2 \pm 0.2\%$
QCD	$5 \times 10^{-5} \pm 6 \times 10^{-5}$	$3 \pm 3\%$

Table 5.1: *Parameters and percentages of the total MC admixture for the different MC samples after pre-selection and MC normalisation.*

Figure 5.10 shows the resulting distributions for the variables used in the fit, $p_{\mathrm{T}}^{\mathrm{jet2}}$ and $M_{\mathrm{J2,J3}}$, and other variables such as M_{JJJ}, M_{JJ}, $p_{\mathrm{T}}^{\mathrm{J1}}/p_{\mathrm{T}}^{\mathrm{jet1}}$ and $\sqrt{\mathrm{dmin}_{21}}$. In the six plots the total MC distribution follows closely the data, and the dominance of W+jets as the main source of background, with a contribution of 68%, is established. In addition the measured M_{JJJ} distribution (figure 5.10c) shows an indication of a peak around M_{top}, which is correctly represented by the MC admixture.

Other variables were also considered for the fit (see section 5.6.2), finding that $p_{\mathrm{T}}^{\mathrm{jet2}}$ and $M_{\mathrm{J2,J3}}$ provides the best performance.

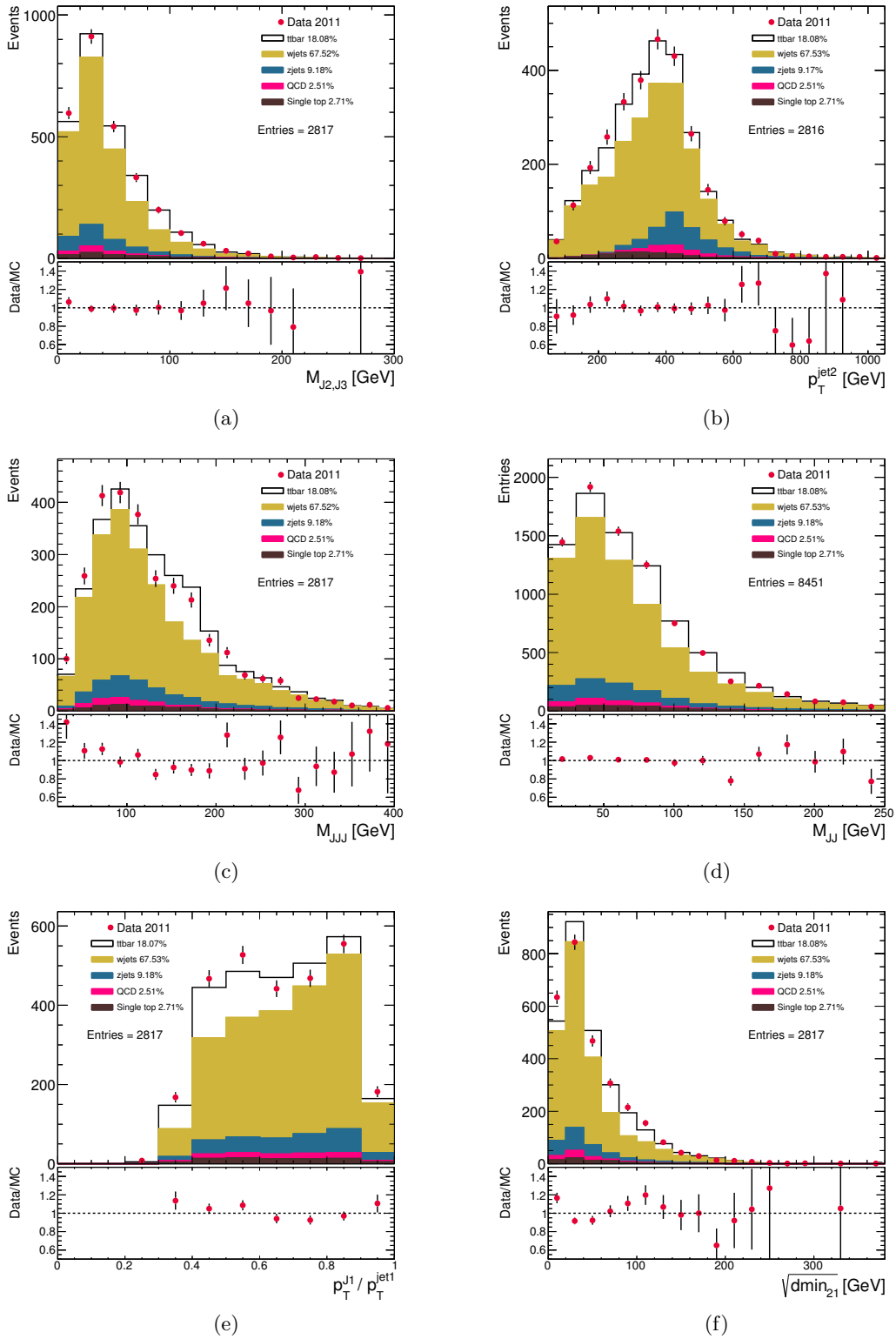


Figure 5.10: *Distributions after pre-selection for data (dots) and MC (histograms) as functions of* $M_{\mathrm{J2,J3}}$ *(a),* $p_{\mathrm{T}}^{\mathrm{jet2}}$ *(b),* M_{JJJ} *(c),* M_{JJ} *(d),* $p_{\mathrm{T}}^{\mathrm{J1}}/p_{\mathrm{T}}^{\mathrm{jet1}}$ *(e) and* $\sqrt{\mathrm{dmin}_{21}}$ *(f). The MC contributions obtained by the MC normalisation procedure are stacked one on top of the other; the black line outer line of the histogram represents the sum of all MC samples mixed according to the results of the fit and each colour represent the individual samples. The fraction of each of the samples in the final admixture is also indicated. The lower part of the plots show the ratio between data and the total MC prediction.*

5.4 Final selection

To improve the background rejection, additional requirements were defined based on the substructure properties of the top candidate:

- events were rejected if the distance in the $\eta - \phi$ plane between the sub-leading small-R jet and the large-R jet ($\Delta R_{\mathrm{J2,jet1}}$) was lower than 0.2;
- events were rejected if the fraction between the p_{T} of the sub-leading small-R jet and the p_{T} of the large-R jet ($p_{\mathrm{T}}^{\mathrm{J2}}/p_{\mathrm{T}}^{\mathrm{jet1}}$) was lower than 0.1;
- events were rejected if the distance in the $\eta - \phi$ plane between the leading and sub-leading small-R jets ($\Delta R_{\mathrm{J2,J1}}$) was lower than 0.3 or higher than 1.1.

Figure 5.11 shows the distributions of the variables used in the final selection, highlighting the differences between the different MC samples and showing (dashed lines) the location of the chosen cuts. In figure 5.12 the effect of the final selection on the $t\bar{t}$ MC samples is shown for $p_{\mathrm{T}}^{\mathrm{jet2}}$, $M_{\mathrm{J2,J3}}$, M_{JJJ}, M_{JJ}, $p_{\mathrm{T}}^{\mathrm{J1}}/p_{\mathrm{T}}^{\mathrm{jet1}}$ and $\sqrt{\mathrm{dmin}_{21}}$. These distributions show that the effect of the final requirements on the $t\bar{t}$ samples is under control and that the shape of the variables does not change dramatically, making them good choices for the analysis.

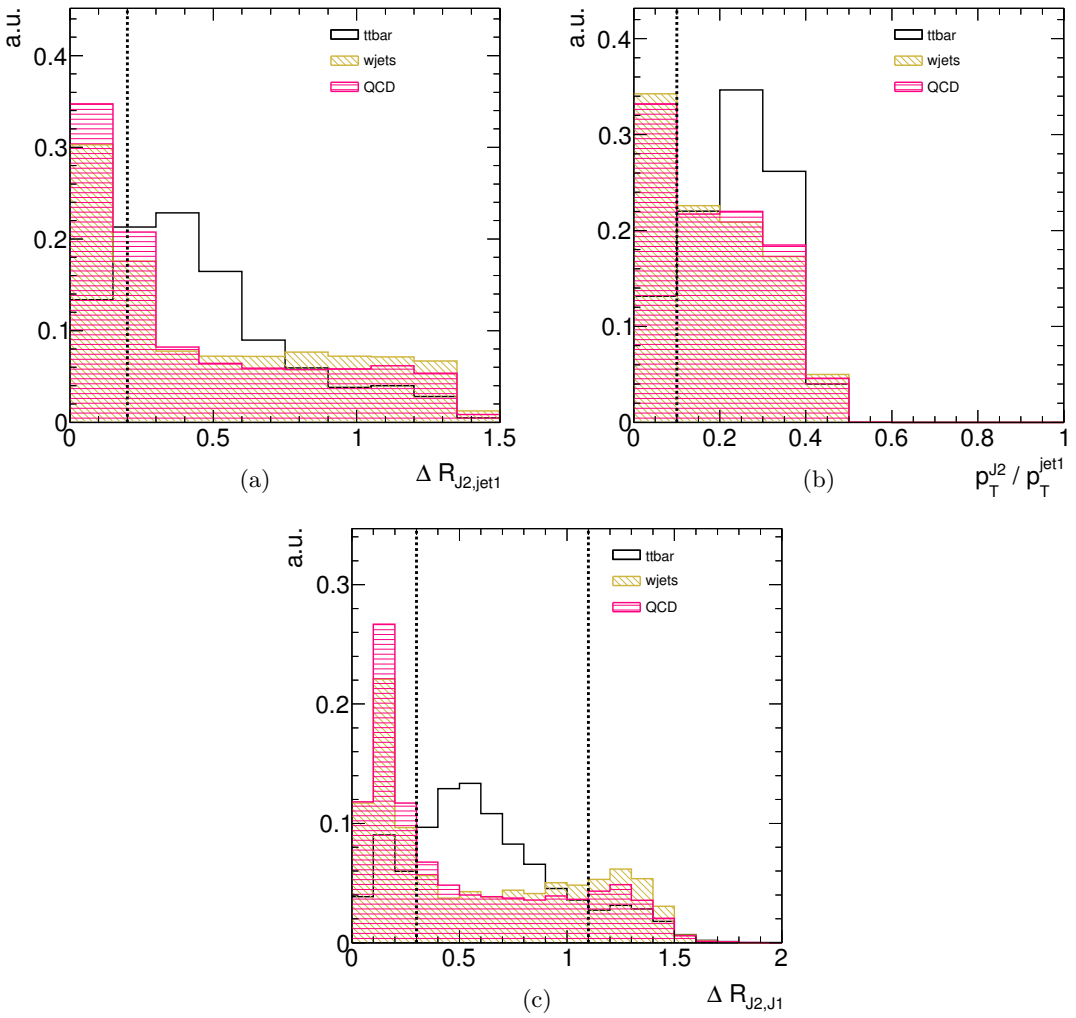


Figure 5.11: *MC distributions after pre-selection for the signal and the most important backgrounds as functions of* $\Delta R_{\mathrm{J2,jet1}}$ *(a),* $p_{\mathrm{T}}^{\mathrm{J2}}/p_{\mathrm{T}}^{\mathrm{jet1}}$ *(b) and* $\Delta R_{\mathrm{J2,J1}}$ *(c). The distributions are normalised to unit area. The dashed lines represent the location of the final selection cuts.*

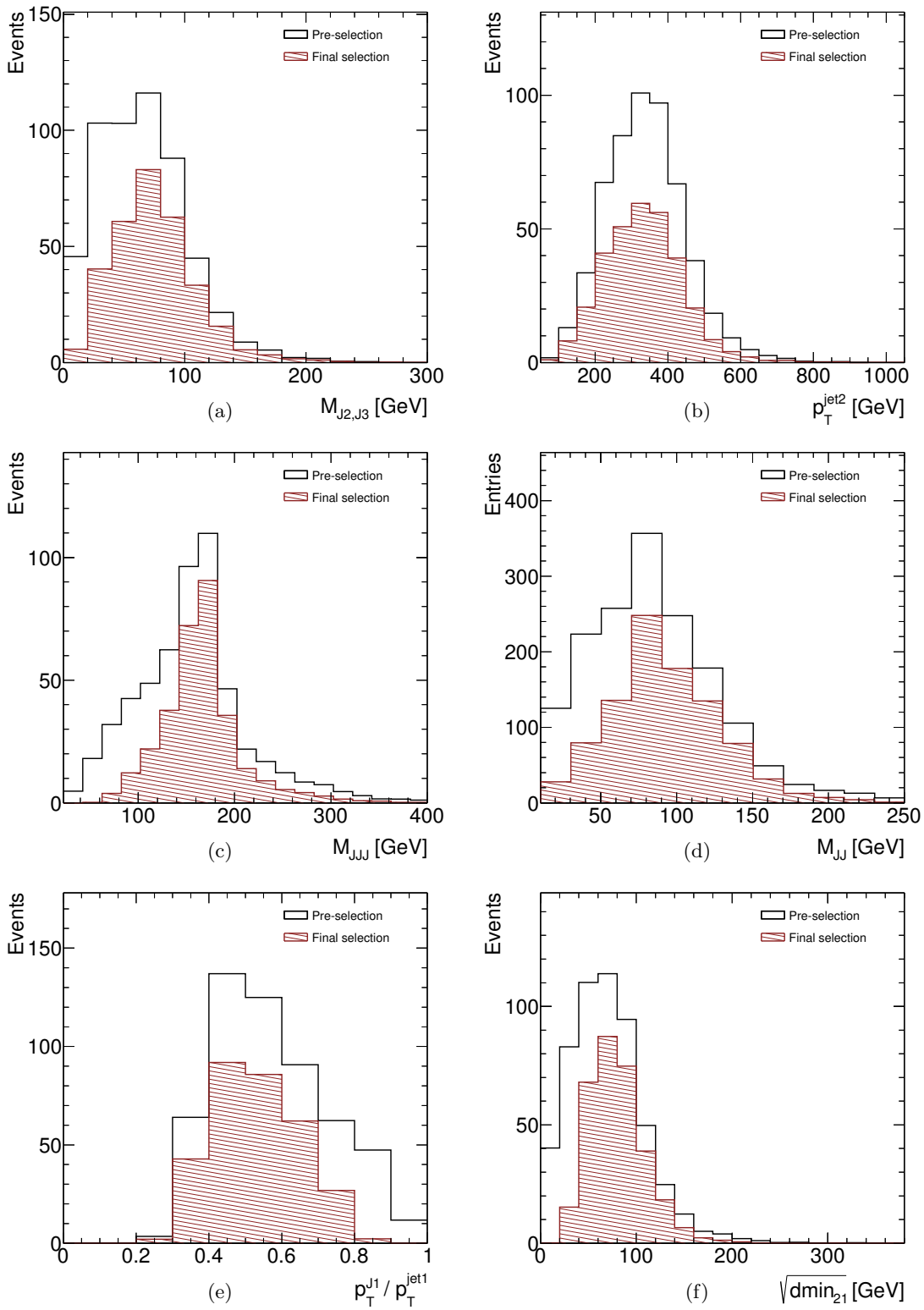


Figure 5.12: *MC distributions after pre-selection and final selection for the $t\bar{t}$ samples as functions of $M_{J2,J3}$ (a), p_T^{jet2} (b), M_{JJJ} (c), M_{JJ} (d), p_T^{J1}/p_T^{jet1} (e) and $\sqrt{dmin_{21}}$ (f).*

5.5 Efficiency

To calculate the global efficiency for this method, a selection that uses information from the particle and parton-levels of the MC simulation (see chapter 3) was defined which follows closely the selection of the semileptonic $t\bar{t}$ events at detector level. These selection criteria, applied over the $t\bar{t}$ samples and referred to as truth selection, are:

- direct decay products of the top quarks were analysed (at the parton-level) and events with only one electron among those products were selected. The electron was required to have $|\eta| < 2.47$ and $p_{\mathrm{T}} > 40$ GeV and events in which the electron lies within the crack region, $1.37 < |\eta| < 1.52$, were rejected;
- stable particles (at the particle level) were clustered using the anti-k_{T} algorithm and $R = 1.5$. Neutrinos and muons were not included;
- events were required to have at least one large-R jet with $p_{\mathrm{T}} > 400$ GeV;
- top candidates were selected using exactly the same procedure as for the detector level, namely, the leading large-R jet and the electron selected as explained above;
- events were rejected in which the distance between the selected electron and the jet closest to the electron in the $\eta - \phi$ plane was less than 0.6. This accounts for the isolation requirement at detector level. For this step, jets reconstructed with the anti-k_{T} algorithm and R=0.4 from the stable particles, excluding the selected electron and imposing a minimum p_{T} of 10 GeV were used;
- The same requirements over the substructure variables as for the detector level were required for the final selection, namely, $\Delta R_{\mathrm{J2,jet1}} > 0.2$, $p_{\mathrm{T}}^{\mathrm{J2}}/p_{\mathrm{T}}^{\mathrm{jet1}} > 0.1$ and $0.3 < \Delta R_{\mathrm{J2,J1}} < 1.1$.

The global efficiency was defined as

$$\epsilon = \frac{N_{truth,reco}}{N_{truth}}, \tag{5.1}$$

where $N_{truth,reco}$ is the number of events that pass both the truth and detector level selections and N_{truth} is the number of events that pass the truth selection. Using this definition and the truth selection described above a global efficiency of approximately 60% was obtained.

Figure 5.13 shows the comparison of the $M_{\mathrm{J2,J3}}$, $p_{\mathrm{T}}^{\mathrm{jet2}}$, M_{JJJ}, M_{JJ}, $p_{\mathrm{T}}^{\mathrm{J1}}/p_{\mathrm{T}}^{\mathrm{jet1}}$ and $\sqrt{\mathrm{dmin}_{21}}$ distributions for the detector level and truth selections when performed over the $t\bar{t}$ MC sample. It is observed that the method does not introduce any critical change in the shape of the detector level distributions and that they closely resemble those obtained at the particle level.

5.5.1 Top-tagging efficiency

The top-tagging efficiency was defined as:

$$\epsilon_{tagging} = \frac{N^{jets}_{matched,tagged}}{N^{jets}_{matched}}, \tag{5.2}$$

where $N^{jets}_{matched,tagged}$ is the number of tagged large-R jets at detector level matched with large-R jets at particle level and $N^{jets}_{matched}$ is the total number of large-R jets at detector level

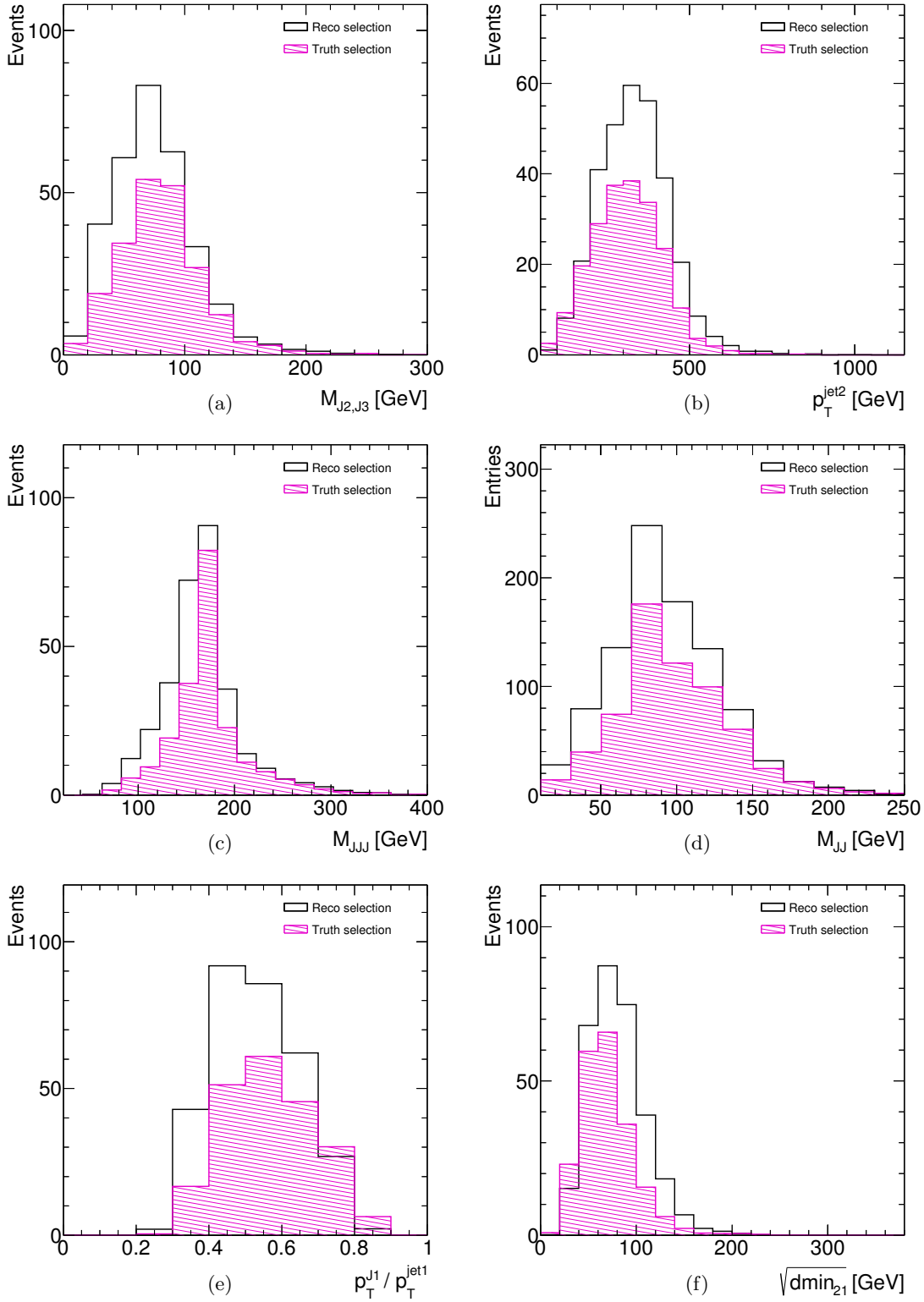


Figure 5.13: *MC distributions after detector level and truth selections for the $t\bar{t}$ samples as functions of $M_{J2,J3}$ (a), p_T^{jet2} (b), M_{JJJ} (c), M_{JJ} (d), p_T^{J1}/p_T^{jet1} (e) and $\sqrt{dmin_{21}}$ (f).*

matched with large-R jets at particle level. Large-R jets were tagged using the substructure requirements defined in section 5.4, i.e. a jet is considered "tagged" if it passes those requirements. Large-R jets at particle level were obtained by clustering stable particles using the anti-k_{T} algorithm with $R = 1.5$; neutrinos and muons were not included.

The number of matched and tagged-and-matched jets were obtained for all the MC samples after pre-selection, considering only the leading large-R jet (or the sub-leading large-R jet if its distance in the $\eta - \phi$ plane to the selected electron was larger than 1.5) for each event. The large-R jet at detector level was considered matched to a large-R jet at particle level if the distance between them in the $\eta - \phi$ plane was less than 1.125 (3/4 of the distance parameter). Large-R jets at particle level were required to have $p_{\mathrm{T}} > 350$ GeV and, in the case of $t\bar{t}$ MC samples, were only considered as candidates for the matching if the distance in the $\eta - \phi$ plane between the large-R jet and the hadronically decaying top was less than 1.125.

Using the samples described above a top-tagging efficiency of approximately 77% was obtained, with rejection factors (defined as $1/\epsilon_{tagging}$) of approximately 5.9 for W+jets, 5.3 for Z+jets, 3.1 for single top and 6.7 for QCD. This tagging efficiencies are similar or better than for other top-taggers currently being used by the ATLAS collaboration [86, 87]

5.6 Results

A sample of boosted $t\bar{t}$ candidates in the semileptonic channel was selected in data by applying the method explained in the previous sections. The number of data events and expected MC events after pre- and final selections are detailed in Table 5.2.

Sample	Pre-selection	Final-selection
$t\bar{t}$	505	336
Single top	76	29
W+jets	1886	364
Z+jets	256	55
QCD	70	12
MC combination	2793	796
Data	**2817 ± 53**	**796 ± 28**

Table 5.2: *Number of data events and expected MC events after pre- and final selections. The final-selection includes a global normalisation of the MC combination to the data.*

Figures 5.14a and 5.14b show the M_{JJJ}, M_{JJ} distributions, respectively, for data and signal and background MC. The M_{JJJ} and M_{JJ} data distributions show a clear peak around M_{top} and M_{W}, respectively. The sum of signal and background MC gives a good description of the data distribution. The signal MC predicts 42.3% of $t\bar{t}$ events. The background is dominated by W+jets processes ($\sim 45\%$); the other contributions to the background are small. Figure 5.14 also shows the p_{T} and rapidity of the four-vectorial sum of the three small-R jets ($p_{\mathrm{T}}^{\mathrm{JJJ}}$ and Y_{JJJ}) while figure 5.15 shows the $M_{\mathrm{J2,J3}}$, $p_{\mathrm{T}}^{\mathrm{jet2}}$, $p_{\mathrm{T}}^{\mathrm{J1}}/p_{\mathrm{T}}^{\mathrm{jet1}}$ and $\sqrt{\mathrm{dmin}_{21}}$ distributions. In all cases, there is a good description of the data by the MC.

Figure 5.16 and figure 5.17 show the same distributions as in figure 5.14 and figure 5.15 after the subtraction of the background has been performed. Even though after the final selection the W+jets is still the dominant background and its fraction in the MC admixture ($\sim 45\%$) is similar to that of the signal ($\sim 42\%$), the precision achieved allows the background to be removed without deforming the shape. This can be observed particularly well in figure 5.16a and figure 5.16b where the M_{JJJ} and M_{JJ} peaks around M_{top} and M_{w}, respectively, are properly preserved through the process. There are approximately 336 events in this background-subtracted data sample. A good description of the data distributions by the signal MC is obtained.

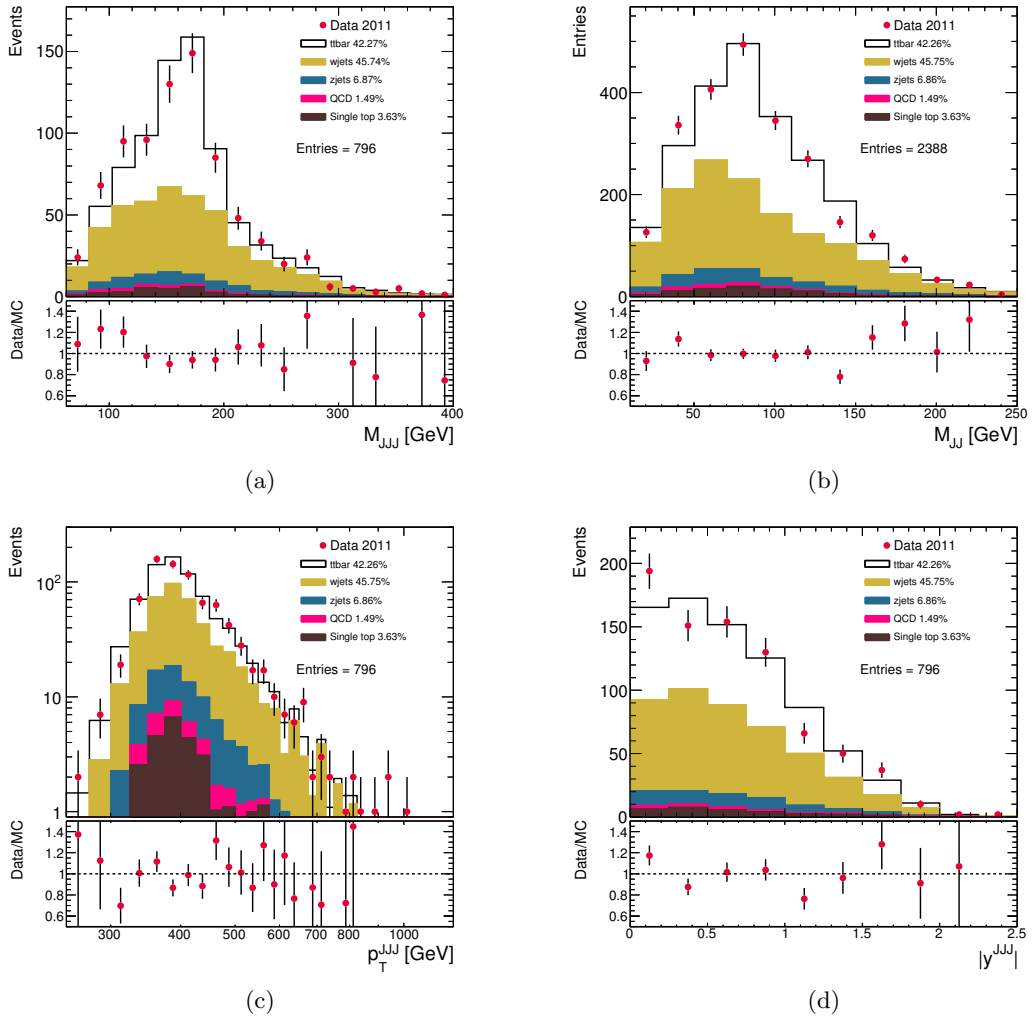


Figure 5.14: *Distributions after the final selection for data (dots) and MC (histograms) as functions of M_{JJJ} (a), M_{JJ} (b), p_T^{JJJ} (c) and $|y^{JJJ}|$ (d). The MC contributions obtained by the MC normalisation procedure are stacked one on top of the other; the black line outer line of the histogram represents the sum of all MC samples mixed according to the results of the fit and each colour represent the individual samples.The fraction of each of the samples in the final admixture is also indicated. The lower part of the plots show the ratio between data and the total MC prediction.*

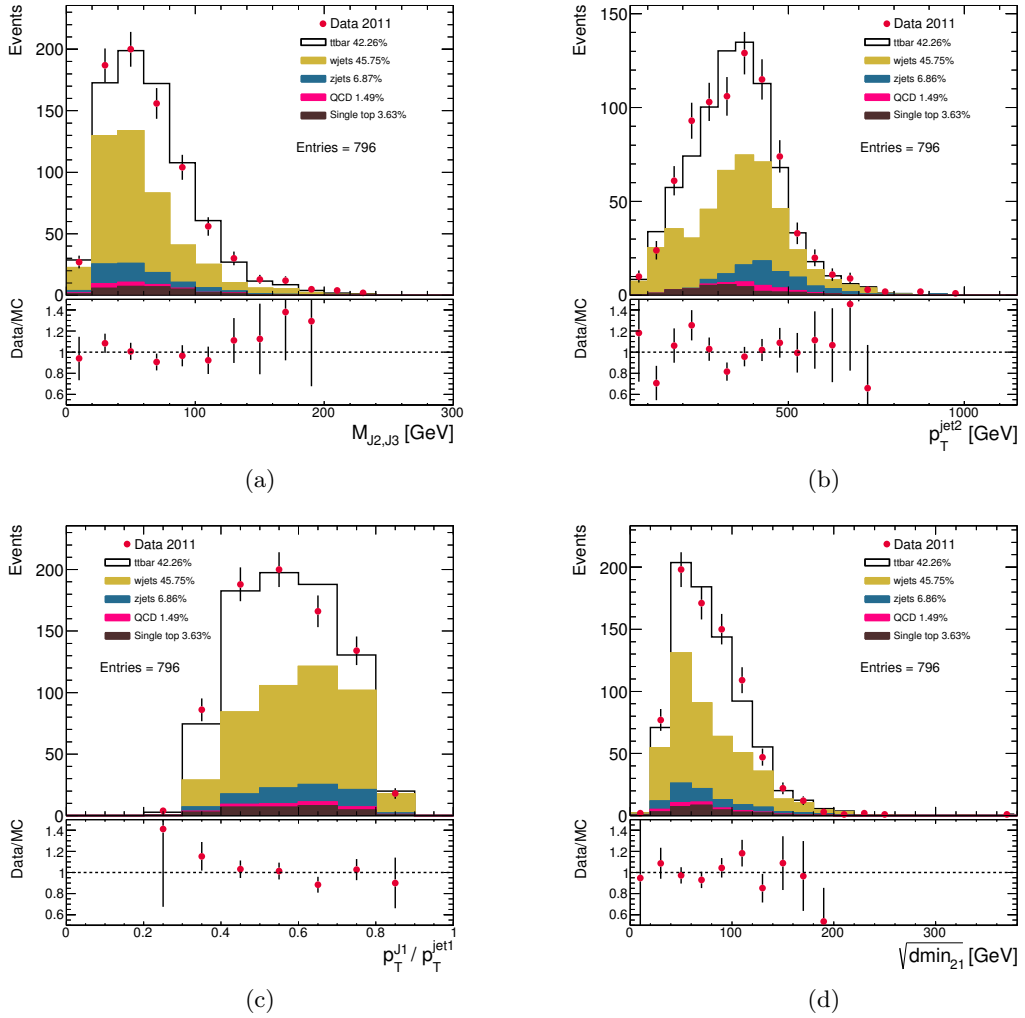


Figure 5.15: *Distributions after the final selection for data (dots) and MC (histograms) as functions of $M_{\mathrm{J2,J3}}$ (a), $p_{\mathrm{T}}^{\mathrm{jet2}}$ (b), $p_{\mathrm{T}}^{\mathrm{J1}}/p_{\mathrm{T}}^{\mathrm{jet1}}$ (c) and $\sqrt{\mathrm{dmin}_{21}}$ (d). The MC contributions obtained by the MC normalisation procedure are stacked one on top of the other; the black line outer line of the histogram represents the sum of all MC samples mixed according to the results of the fit and each colour represent the individual samples. The fraction of each of the samples in the final admixture is also indicated. The lower part of the plots show the ratio between data and the total MC prediction.*

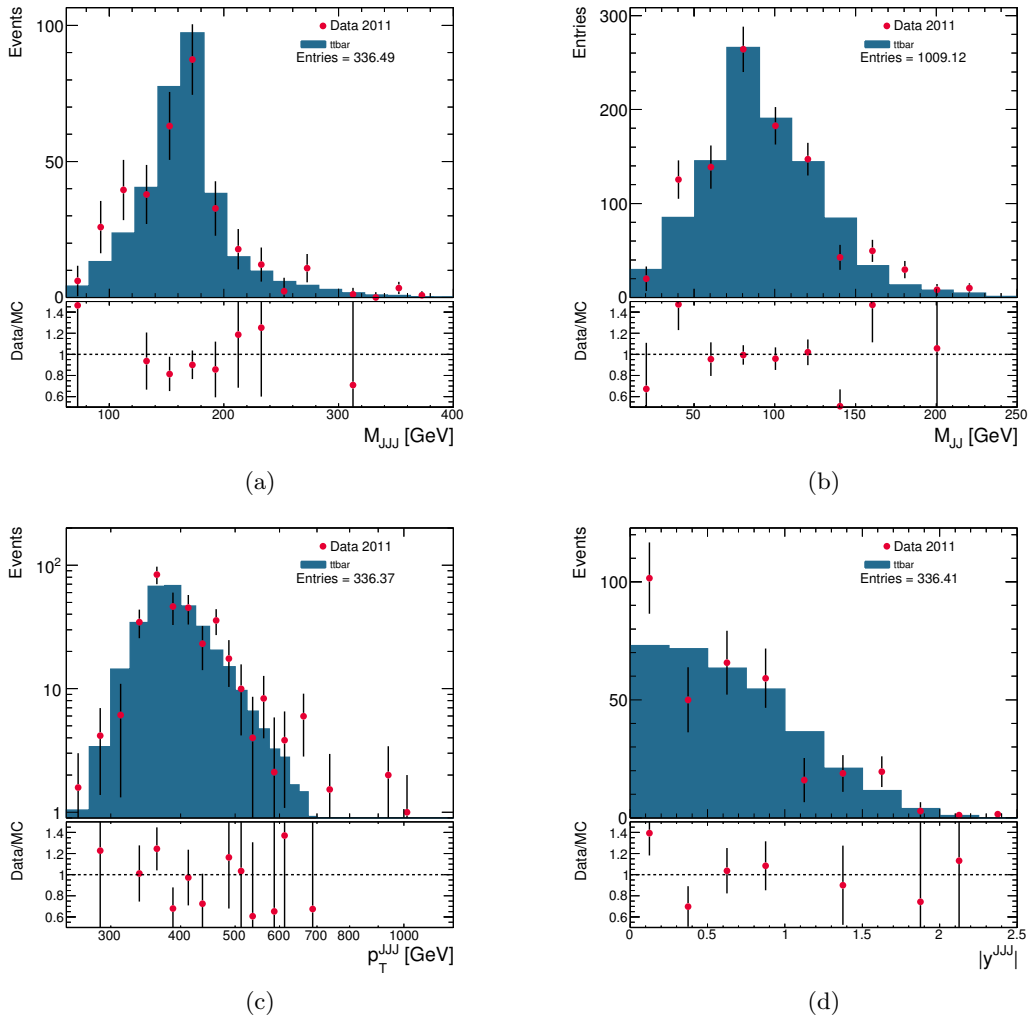


Figure 5.16: *Distributions after background subtraction for data (dots) and the $t\bar{t}$ MC sample (histograms) as functions of M_{JJJ} (a), M_{JJ} (b), $p_{\mathrm{T}}^{\mathrm{JJJ}}$ (c) and $|y^{\mathrm{JJJ}}|$ (d). The lower part of the plots show the ratio between data and the $t\bar{t}$ MC sample prediction.*

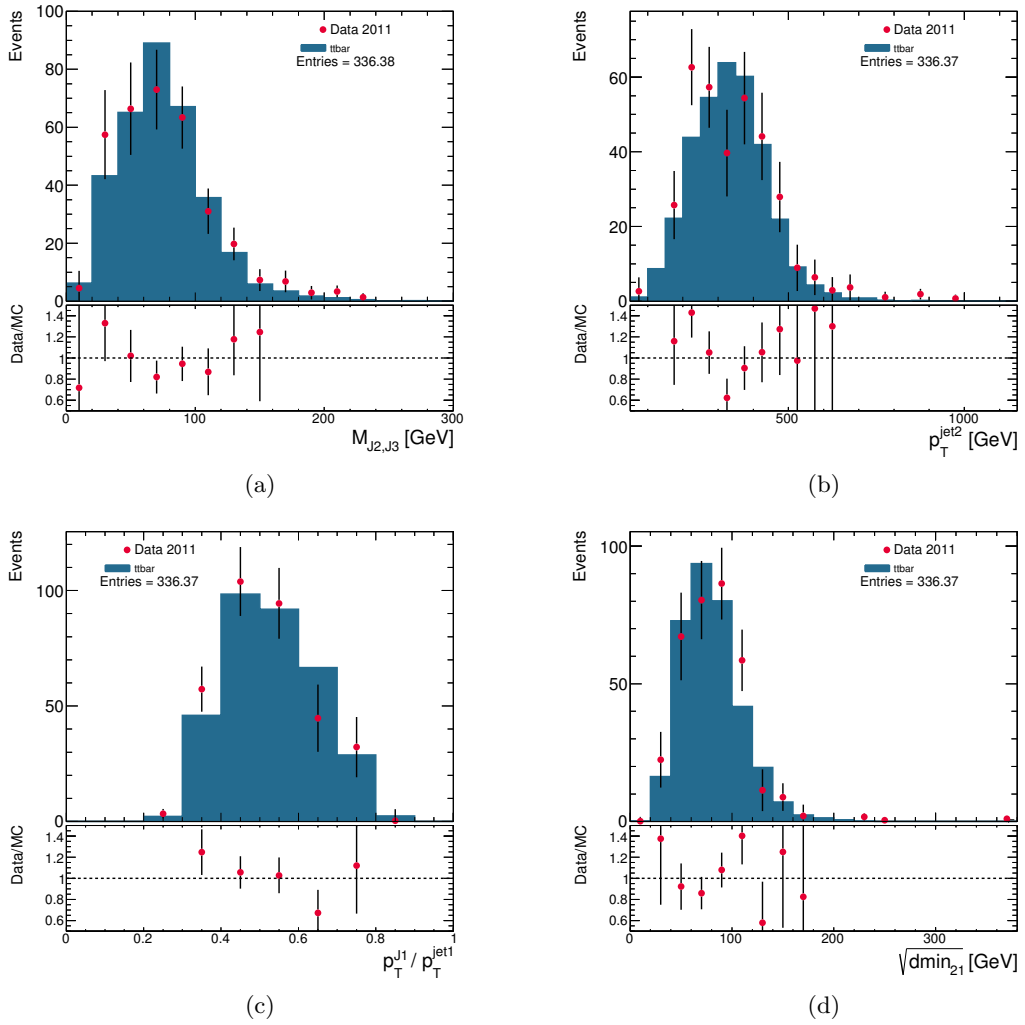


Figure 5.17: *Distributions after background subtraction for data (dots) and the $t\bar{t}$ MC sample (histograms) as functions of $M_{\mathrm{J2,J3}}$ (a), $p_{\mathrm{T}}^{\mathrm{jet2}}$ (b), $p_{\mathrm{T}}^{\mathrm{J1}}/p_{\mathrm{T}}^{\mathrm{jet1}}$ (c) and $\sqrt{\mathrm{dmin}_{21}}$ (d). The lower part of the plots show the ratio between data and the $t\bar{t}$ MC sample prediction.*

5.6.1 Cross check using b-tagging

A compelling cross check of the method was performed by applying a b-tagging algorithm to the final selected sample. The b-tagging procedure, which was applied both for data and MC samples on top of the final selection, consists of:

- jets reconstructed using the anti-k_{T} algorithm with $R = 0.4$ and calibrated using the EM+JES scheme were used as the starting point;
- jets were tagged using the MV1 algorithm (see section 4.5.6). A jet was required to have a mv1 weight higher than 0.601713 to be considered as b-tagged, which corresponds to a b-tagging efficiency of 70% [83];
- events were required to have at least one b-tagged jet with a distance to the selected electron in the $\eta - \phi$ plane less than 1.5. The distance requirement was implemented to further reduce the W+jets background: if the W boson has decayed into an electron and a neutrino ($W \to e\nu$), which is heavily enhanced by the event selection, the b-quark has to come from the extra partons, which are expected to be well separated from the W decays products.

For the MC samples, scale factors were applied to compensate for the differences in efficiency between data and MC. In the case of the QCD MC samples, a higher efficiency working point was used ($mv1 > 0.0714225$ corresponding to a 85% efficiency).

The background-subtracted data sample contains 188 events, which corresponds to 56% of the nominal sample. Figure 5.18 shows the distributions after b-tagging as functions of $p_{\mathrm{T}}^{\mathrm{jet2}}$, $M_{\mathrm{J2,J3}}$, M_{JJJ}, M_{JJ}, $p_{\mathrm{T}}^{\mathrm{J1}}/p_{\mathrm{T}}^{\mathrm{jet1}}$ and $\sqrt{\mathrm{dmin}_{21}}$ for data and MC. The distributions show the same characteristics as in figure 5.14 and figure 5.15 while at the same time showing the reduction of background that the b-tagging procedure provides.

Figure 5.19 shows the normalised data distributions before and after b-tagging (after subtracting the background in both cases) for the M_{JJJ} and M_{JJ} distributions. It is observed that the shape of the distributions is the same, which supports the assumption that the nominal selection is able to properly select $t\bar{t}$ candidates.

5.6.2 Use of alternative variables in the fit

Section 5.3 contains what was found to be the optimal pair of variables to do the fit. However, other sets of variables were also tested. Figure 5.20 shows the normalised distributions for one of those sets: the distance in the $\eta - \phi$ plane between the sub-sub-leading small$-R$ jet and the large$-R$ jet ($\Delta R_{\mathrm{J3,jet1}}$) and $p_{\mathrm{T}}^{\mathrm{jet2}}$. Figure 5.21 shows that the correlation between both variables is not strong.

The same procedure described in section 5.3 was also applied, albeit using the alternative set of variables, which in turn produced another set of parameters (α, β and γ) and therefore different fractions for each MC contribution; the results are compared to the nominal ones in table 5.3. The reduced χ^2 obtained was 1.29, slightly bigger than in the nominal case. Distributions for $p_{\mathrm{T}}^{\mathrm{jet2}}$, $\Delta R_{\mathrm{J3,jet1}}$, M_{JJJ}, M_{JJ}, $p_{\mathrm{T}}^{\mathrm{J1}}/p_{\mathrm{T}}^{\mathrm{jet1}}$ and $\sqrt{\mathrm{dmin}_{21}}$ after pre-selection are shown in figure 5.22 and the same distributions after the final selection are shown in figure 5.23. Both figures are very similar to those obtained with the nominal fit variables; this indicates that the selection of the particular set of variables for the fit is not critical to the method as long as they are able to discriminate between the signal and the different background processes.

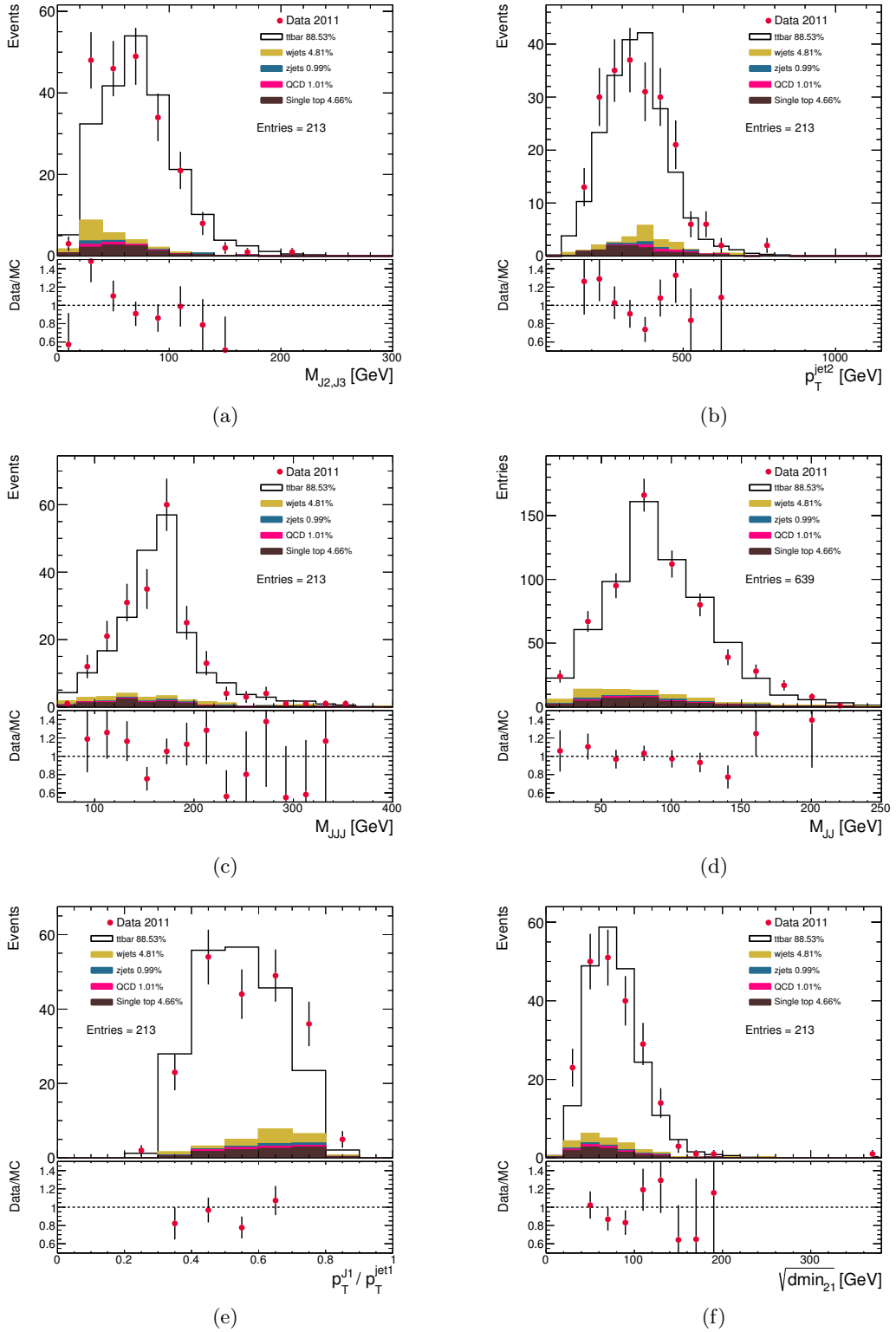
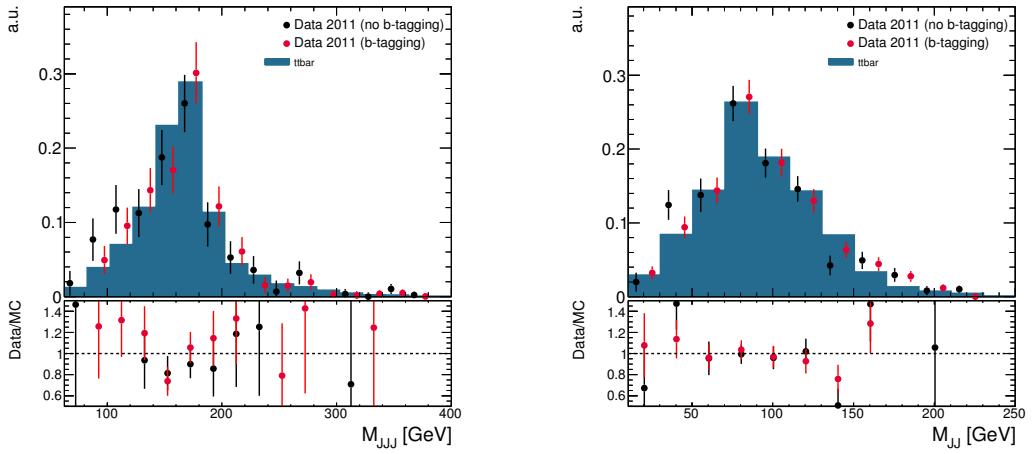
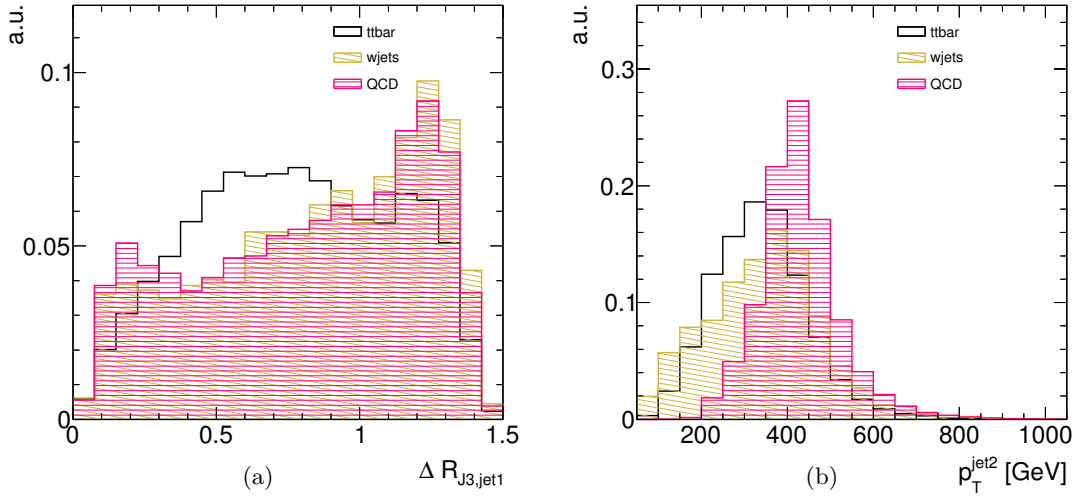


Figure 5.18: *Distributions after the final selection and b-tagging criteria for data (dots) and MC (histograms) as functions of $M_{J2,J3}$ (a), p_T^{jet2} (b) M_{JJJ} (c), M_{JJ} (d), p_T^{J1}/p_T^{jet1} (e) and $\sqrt{dmin_{21}}$. The MC contributions obtained by the MC normalisation procedure are stacked one on top of the other; the black line outer line of the histogram represents the sum of all MC samples mixed according to the results of the fit and each colour represent the individual samples.The fraction of each of the samples in the final admixture is also indicated. The lower part of the plots show the ratio between data and the total MC prediction.*



(a) (b)

Figure 5.19: *Distributions after background subtraction for data without (black dots) and with (red dots) b-tagging and for the $t\bar{t}$ MC sample (histograms) as functions of M_{JJJ} (a) and M_{JJ} (b). All the distributions are normalised to unit area. The lower part of the plots show the ratio between data and the $t\bar{t}$ MC sample prediction.*



(a) (b)

Figure 5.20: *MC distributions after pre-selection for the signal and the most important backgrounds as functions of $\Delta R_{\mathrm{J3,jet1}}$ (a) and $p_{\mathrm{T}}^{\mathrm{jet2}}$ (b). The distributions are normalised to unit area.*

Sample	Nominal parameter	Nominal percentage	alternative parameter	alternative percentage
$t\bar{t}$	0.9 ± 0.1	$18 \pm 2\%$	1.1 ± 0.2	$22 \pm 3\%$
Single top	0.9 ± 0.1	$2.7 \pm 0.3\%$	1.1 ± 0.2	$3.2 \pm 0.5\%$
W+jets	0.94 ± 0.06	$68 \pm 1\%$	0.86 ± 0.07	$62 \pm 2\%$
Z+jets	0.94 ± 0.06	$9.2 \pm 0.2\%$	0.86 ± 0.07	$8.5 \pm 0.3\%$
QCD	$5 \times 10^{-5} \pm 6 \times 10^{-5}$	$3 \pm 3\%$	$9 \times 10^{-5} \pm 6 \times 10^{-5}$	$4 \pm 4\%$

Table 5.3: *Comparison between the parameters and percentages of the final MC admixture for the different MC samples obtained with the nominal and alternative fits.*

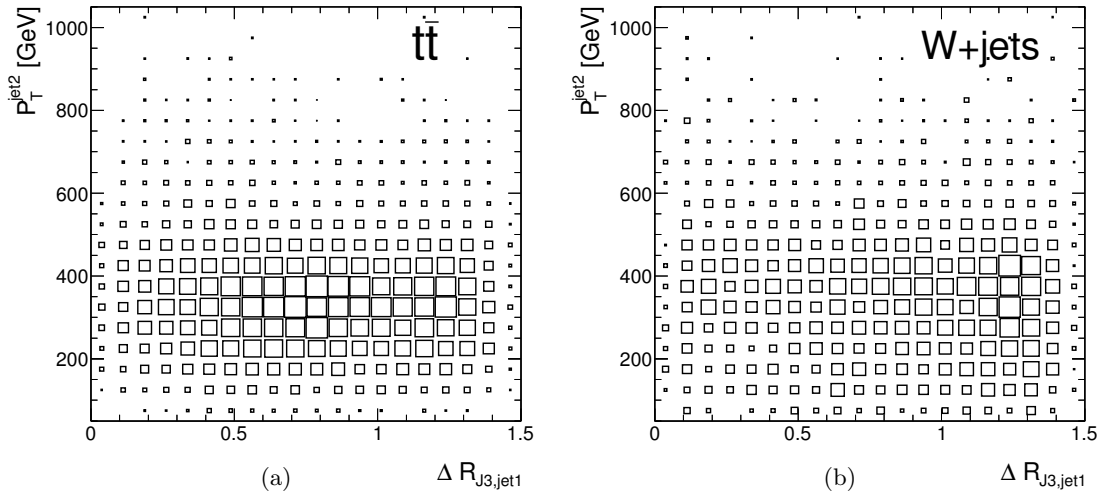


Figure 5.21: *Correlation between* $\Delta R_{\mathrm{J3,jet1}}$ *and* $p_{\mathrm{T}}^{\mathrm{jet2}}$ *for* $t\bar{t}$ *MC samples (a) and* W*+jets MC samples (b).*

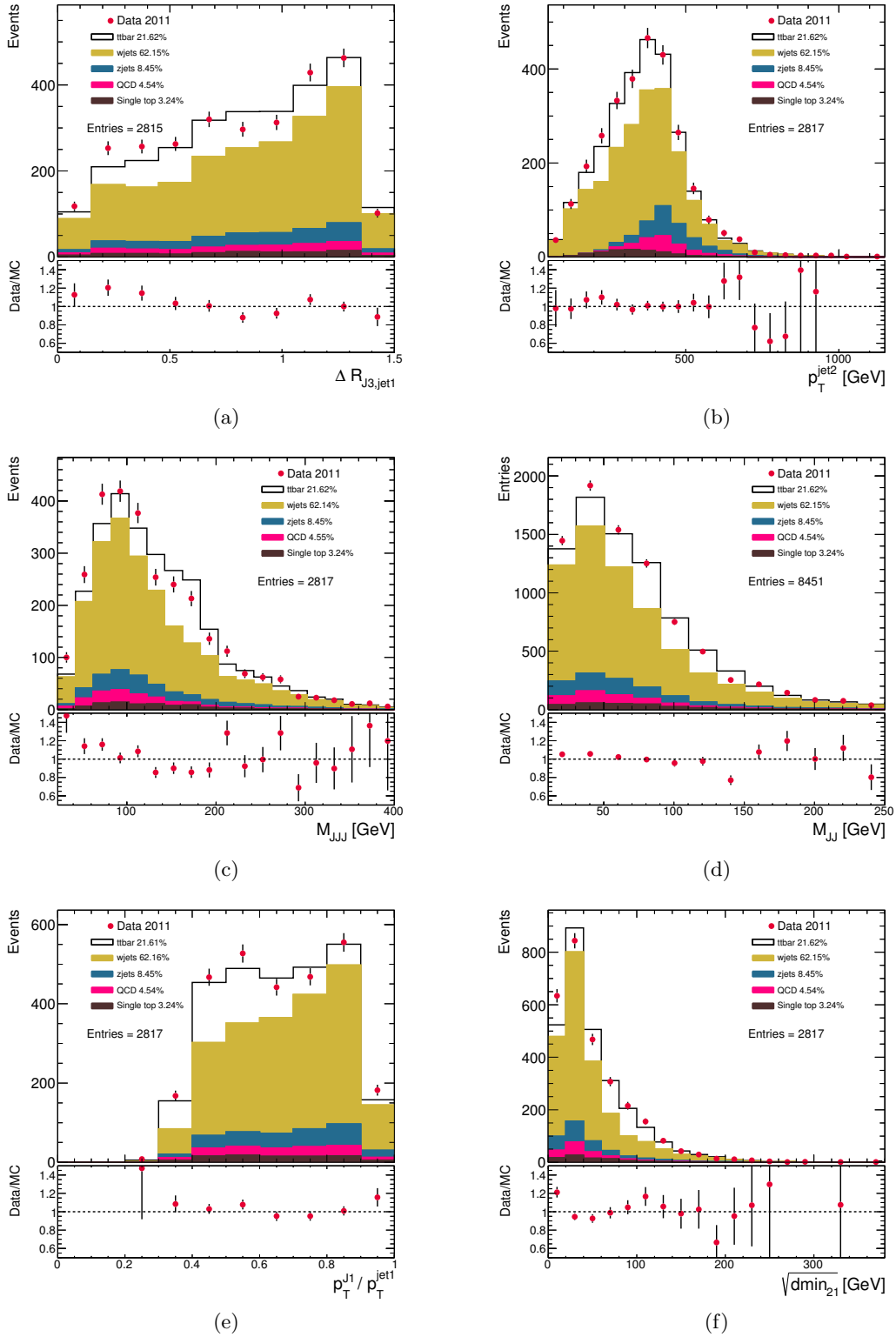


Figure 5.22: *Distributions after pre-selection for data (dots) and MC (histograms) as functions of* $\Delta R_{\mathrm{J3,jet1}}$ *(a),* $p_{\mathrm{T}}^{\mathrm{jet2}}$ *(b),* M_{JJJ} *(c),* M_{JJ} *(d),* $p_{\mathrm{T}}^{\mathrm{J1}}/p_{\mathrm{T}}^{\mathrm{jet1}}$ *(e) and* $\sqrt{\mathrm{dmin}_{21}}$ *(f). The MC contributions obtained by the MC normalisation procedure are stacked one on top of the other; the black line outer line of the histogram represents the sum of all MC samples mixed according to the results of the fit performed using the alternative fit variables and each colour represent the individual samples. The fraction of each of the samples in the final admixture is also indicated. The lower part of the plots show the ratio between data and the total MC prediction.*

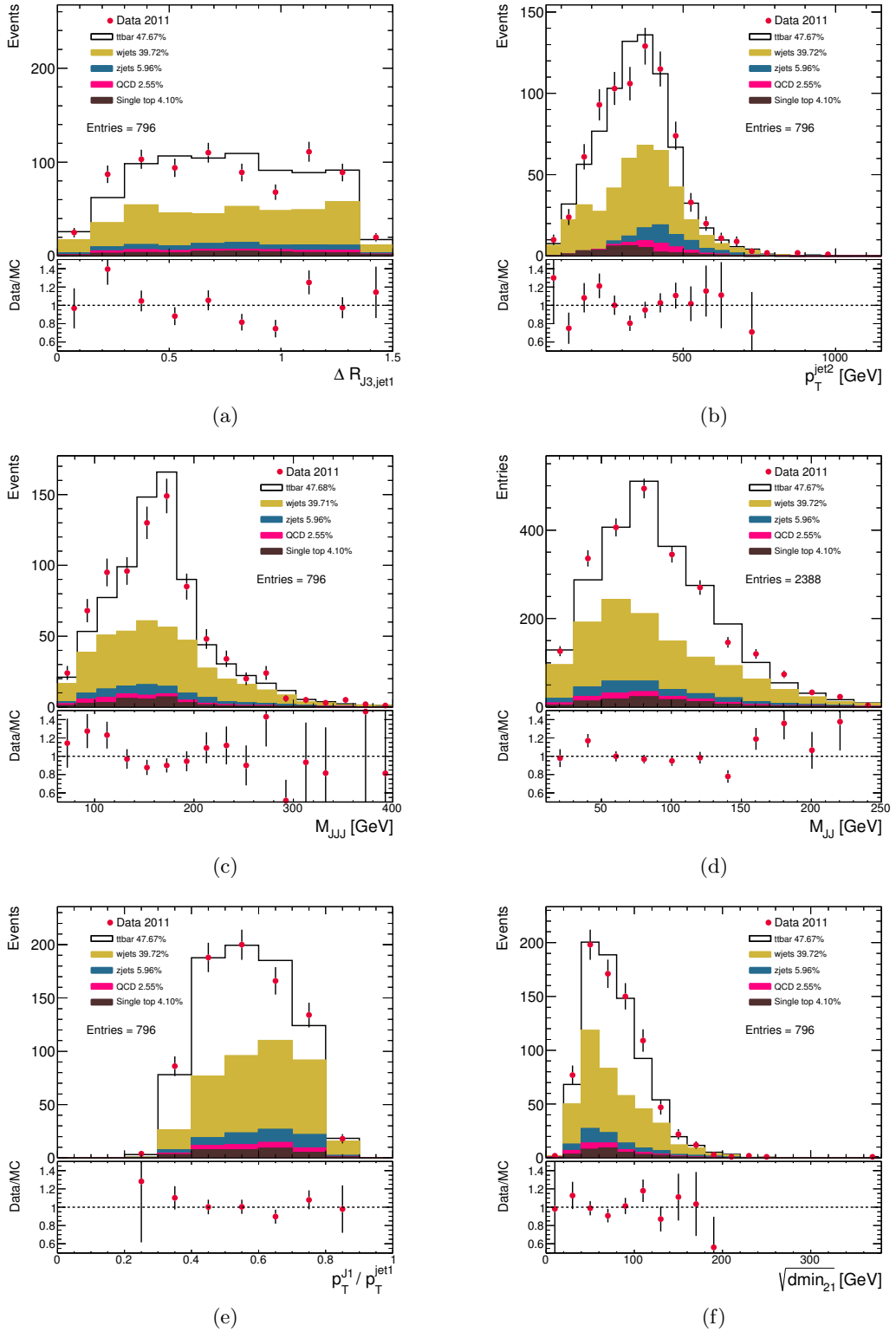


Figure 5.23: *Distributions after final selection for data (dots) and MC (histograms) as functions of* $\Delta R_{\mathrm{J3,jet1}}$ *(a),* $p_{\mathrm{T}}^{\mathrm{jet2}}$ *(b),* M_{JJJ} *(c),* M_{JJ} *(d),* $p_{\mathrm{T}}^{\mathrm{J1}}/p_{\mathrm{T}}^{\mathrm{jet1}}$ *(e) and* $\sqrt{\mathrm{dmin}_{21}}$ *(f). The MC contributions obtained by the MC normalisation procedure are stacked one on top of the other; the black line outer line of the histogram represents the sum of all MC samples mixed according to the results of the fit performed using the alternative fit variables and each colour represent the individual samples. The fraction of each of the samples in the final admixture is also indicated. The lower part of the plots show the ratio between data and the total MC prediction.*

Chapter 6

High-E_T isolated-photon plus jets production at $\sqrt{s} = 8$ TeV with the ATLAS detector

In this chapter, a study of the dynamics of isolated-photon plus one-, two- and three-jet production in pp collisions at a centre-of-mass energy of 8 TeV with the ATLAS detector is presented. Measurements of isolated-photon plus jets differential cross sections are presented as functions of E_{T}^{γ} and $p_{\mathrm{T}}^{\mathrm{jet}}$. The differential cross sections as functions of the difference in azimuthal angle between the photon and the jet, the difference in azimuthal angle between the jets, the photon-jet invariant mass and the scattering angle in the photon-jet centre-of-mass system are presented. Colour-coherence effects were investigated in events with a photon accompanied by two jets by measuring the distributions of the angle of the sub-leading jet around the photon, β^{γ}, or the leading jet, $\beta_{\mathrm{lead}}^{\mathrm{jet}}$, in the $\eta - \phi$ plane; next-to-leading-order QCD calculations are compared to the measurements when available.

6.1 MC simulations

Samples of MC events were generated to study the characteristics of signal and background. The MC samples were also used to determine the response of the detector to jets of hadrons and the correction factors necessary to obtain the particle level cross sections. In addition, samples were also used to estimate hadronisation corrections to the NLO QCD calculations. The samples used in this analysis are detailed in section 3.2.2.

6.2 Data selection

The data used in this analysis were collected during the 2012 LHC running period at a centre-of-mass energy of 8 TeV (see section 2.2). Only good quality data in which every part of the ATLAS detector was fully operation were used. This data set corresponds to an integrated luminosity of 20.28 ± 0.57 fb^{-1} [88]. This luminosity corresponds to the sample collected with the "EF_g120_loose" trigger (see below).

6.2.1 Trigger requirements

The data sample used consists of events triggered by the EF_g120_loose chain, which consists of a high-level photon trigger with a transverse energy threshold of 120 GeV seeded by

the L1_EM30 trigger with a 30 GeV threshold. It contains the loose-level trigger photon identification criteria. See section 4.1 for details on the photon triggers. The EF_g120_loose chain was the lowest-threshold unprescaled photon trigger during the 2012 data-taking period. The trigger requirement was not applied on the MC samples.

The selection criteria applied by the trigger are looser than the photon identification criteria applied in the offline analysis (see section 4.4.2) and allow a plateau of constant efficiency close to 100%. An study on the efficiency of the trigger is included in section 7.2.

6.3 Event selection

The sample of isolated-photon plus jets events was selected using the following criteria:

- events were required to have at least one reconstructed primary vertex, with at least two associated tracks with a minimum p_{T} of 500 MeV, consistent with the average beam-spot position;
- photon candidates were reconstructed following the guidelines in section 4.4.1 and selected using the following criteria:
  - candidates reconstructed near regions of the calorimeter affected by calorimeter failures were not considered;
  - the candidate energy was calibrated using the procedure described in section 4.4.4. A smearing in energy was applied to the reconstructed photon candidates in MC simulated events instead.
  - candidates were required to pass the *loose prime* identification criteria (see section 4.4.2);
  - in events with multiple candidates satisfying these requirements, the candidate with highest transverse energy (leading photon) was retained for further study;
- events in which the leading photon had transverse energy $E_T^{\gamma} > 130$ GeV and $|\eta^{\gamma}| < 2.37$ were selected. The event was excluded if $1.37 < |\eta^{\gamma}| < 1.56$. In addition, the leading photon was required to pass the following criteria:
  - the leading photon was required to pass the *tight* identification criteria (see section 4.4.2);
  - the leading photon was required to be isolated (see section 4.4.3). E_T^{iso} was required to be lower than 4.8 GeV $+4.2 \cdot 10^{-3} \cdot E_T^{\gamma}$, which amounts to a cut on E_T^{iso} of 5.3(9) GeV for $E_T^{\gamma} = 130(1000)$ GeV. A detailed study on the photon isolation is included in section 7.1;
- jets were reconstructed using the anti-k_{T} algorithm with a distance parameter $R = 0.6$ (see section 4.5.1) and calibrated using the LCW+GSC scheme (see section 4.5.3). They were selected using the following criteria:
  - jets with negative calibrated energy or fulfilling the *looser* identification criteria (see section 4.5.2) were rejected;
  - jets close to the leading photon were discarded (distance in the $\eta - \phi$ plane between them smaller than 1.0);
  - jets with $|y^{\mathrm{jet}}| > 4.4$ were discarded.
- events with at least one jet candidate of calibrated $p_{\mathrm{T}}^{\mathrm{jet}} > 50$ GeV were selected;

- additional corrections were applied to MC events to match the overall event conditions of the data sample and to account for known differences between the data and simulations. They are described in section 6.5;

6.3.1 Event selection at particle level

The selection at particle level (see chapter 3) which was used throughout this analysis uses the following criteria:

- the leading prompt photon was selected and required to have $E_T^\gamma > 130$ GeV and $|\eta^\gamma| < 2.37$. The event was excluded if $1.37 < |\eta^\gamma| < 1.56$.
- $E_{\mathrm{T}}^{\mathrm{iso}}$ for the selected photon is required to be less than 10 GeV. This value of $E_{\mathrm{T}}^{\mathrm{iso}}$ at particle level was obtained from the study included in section 7.1;
- jets were reconstructed with the anti-k_{T} algorithm with distance parameter $R = 0.6$ using all stable particles, only jets with $|y^{\mathrm{jet}}| < 4.4$ were considered;
- jets close to the selected photon were discarded (distance in the $\eta - \phi$ plane between them smaller than 1.0);
- events were required to have at least one jet with $p_{\mathrm{T}}^{\mathrm{jet}} > 50$ GeV.

6.4 Data samples

After the event selection, the 2822703 events selected are further divided into different samples, which were built to study observables sensitive to the dynamics of these processes. The samples are not exclusive, i.e. an event can be part of several samples. The same definitions are used for the detector and particle levels.

The different samples and the observables studied in each sample are described in this section.

6.4.1 Photon+one-jet

In the photon+one-jet sample, the jet with highest transverse momentum (leading jet) was required to have $p_{\mathrm{T}}^{\mathrm{jet}} > 100$ GeV. The number of events selected in this sample amounts to 2451236. Figure 6.1 shows an event selected in this sample. The presence of a photon (red line) and one jet (semi-transparent white regions) is seen.

Figure 6.2 shows the transverse energy of the leading photon (E_{T}^γ), the pseudorapidity of the leading photon ($|\eta^\gamma|$), the azimuthal angle of the leading photon (ϕ^γ) the transverse-momentum of the leading jet ($p_{\mathrm{T,lead}}^{\mathrm{jet}}$), the rapidity of the leading jet ($y_{\mathrm{lead}}^{\mathrm{jet}}$) and the azimuthal angle of the leading jet ($\phi_{\mathrm{lead}}^{\mathrm{jet}}$) distributions. Figure 6.3 shows the distributions for the difference in pseudorapidity between the leading photon and the leading jet ($\Delta\eta^{\gamma-\mathrm{jet\,lead}}$) and the difference in the azimuthal angle between the leading photon and the leading jet ($\Delta\phi^{\gamma-\mathrm{jet\,lead}}$).

6.4.2 Photon+one-jet $m^{\gamma j}$ - $\cos\theta^{\gamma j}$

In addition to the observables described in the photon+one jet sample, the dynamics of the photon plus one jet system was studied by measuring the photon-jet invariant mass ($m^{\gamma j}$) and $\cos\theta^*$ (see section 1.5.3), where the variable θ^* will be referred to as $\theta^{\gamma j}$ henceforth.

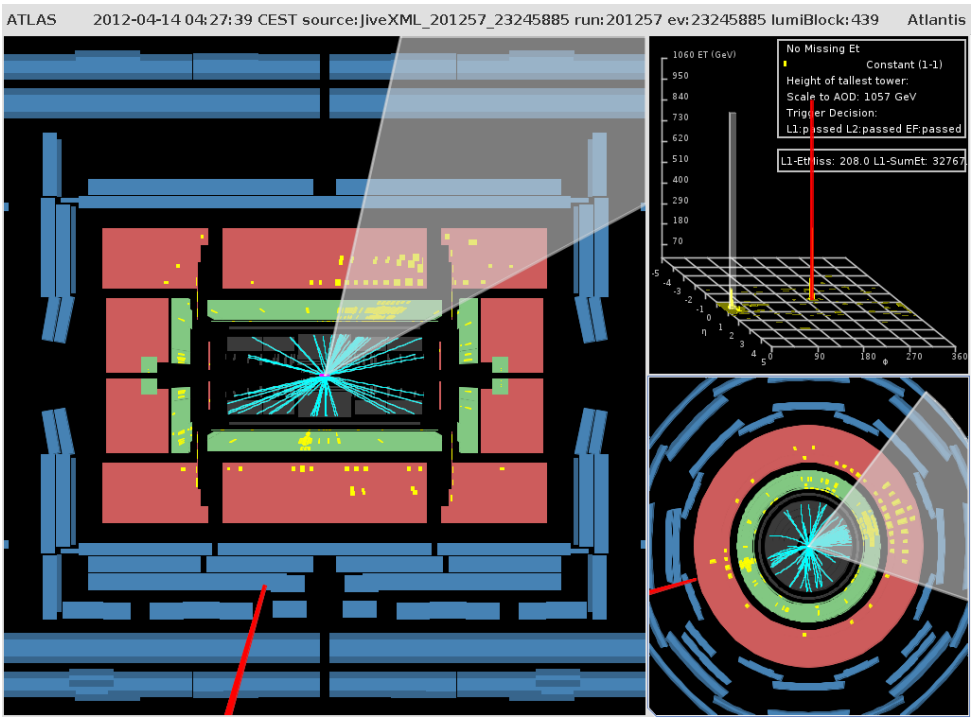


Figure 6.1: *Event display showing an event from the 2012 $\sqrt{s} = 8$ TeV dataset belonging to the photon+ one-jet sample. Red lines represent isolated photons of tight quality, light blue lines represent tracks in the inner detector and yellow squares represent energy deposits in the calorimeters. Jets are represented as semi-transparent white regions. The EM calorimeter is drawn in green while the hadronic calorimeter is drawn in red.*

To study those two observables, additional requirements were imposed to remove the bias due to the rapidity and transverse-momentum cuts on the photon and the jet. The effect of these biases is shown in figure 6.4. To perform unbiased measurements for $m^{\gamma j}$ and $|\cos\theta^{\gamma j}|$ the requirements $|\eta^{\gamma} + y^{\text{jet}}_{\text{lead}}| < 2.37$, $|\cos\theta^{\gamma j}| < 0.83$ and $m^{\gamma j} > 467$ GeV were imposed on top of the photon + jet sample to build the "photon+one-jet $m^{\gamma j}$ - $\cos\theta^{\gamma j}$" sample. The first two requirements avoid the bias induced by the cuts on η^{γ} and $y^{\text{jet}}_{\text{lead}}$, yielding slices of $\cos\theta^{\gamma j}$ with the same length along the $\eta^{\gamma} + y^{\text{jet}}_{\text{lead}}$ axis. The third requirement avoids the bias due to the $E^{\gamma}_{T} > 130$ GeV cut, which can be seen in figure 6.4b. The unbiased kinematic regions are shown as hatched areas in figure 6.4. The number of events selected in the data after these additional requirements is 344572. Figure 6.5 shows the $m^{\gamma j}$ and $|\cos\theta^{\gamma j}|$ distributions.

6.4.3 Photon+two-jets

Photon+two-jets events were studied by selecting events in which, in addition to the photon+ one-jet requirements (see section 6.4.1), there is a second jet (sub-leading jet) with $p^{\text{jet}}_{\text{T}} > 65$ GeV. The number of data events selected in this sample amounts to 567796. Figure 6.6 shows an event selected in this sample. The presence of a photon (red line) and two jets (semi-transparent white regions) is seen.

Figure 6.7 shows the E^{γ}_{T}, $|\eta^{\gamma}|$, ϕ^{γ}, the transverse-momentum of the sub-leading jet ($p^{\text{jet}}_{\text{T,sublead}}$), the rapidity of the sub-leading jet ($|y^{\text{jet}}_{\text{sublead}}|$) and the azimuthal angle of the sub-leading jet ($\phi^{\text{jet}}_{\text{sublead}}$) distributions. Figure 6.8 shows the distributions for the pseudorapidity and azimuthal angle differences between the photon and the sub-leading jet ($\Delta\eta^{\gamma-\text{jet sublead}}$ and $\Delta\phi^{\gamma-\text{jet sublead}}$) and between the leading jet and the sub-leading jet ($\Delta\eta^{\text{jet lead}-\text{jet sublead}}$ and $\Delta\phi^{\text{jet lead}-\text{jet sublead}}$)

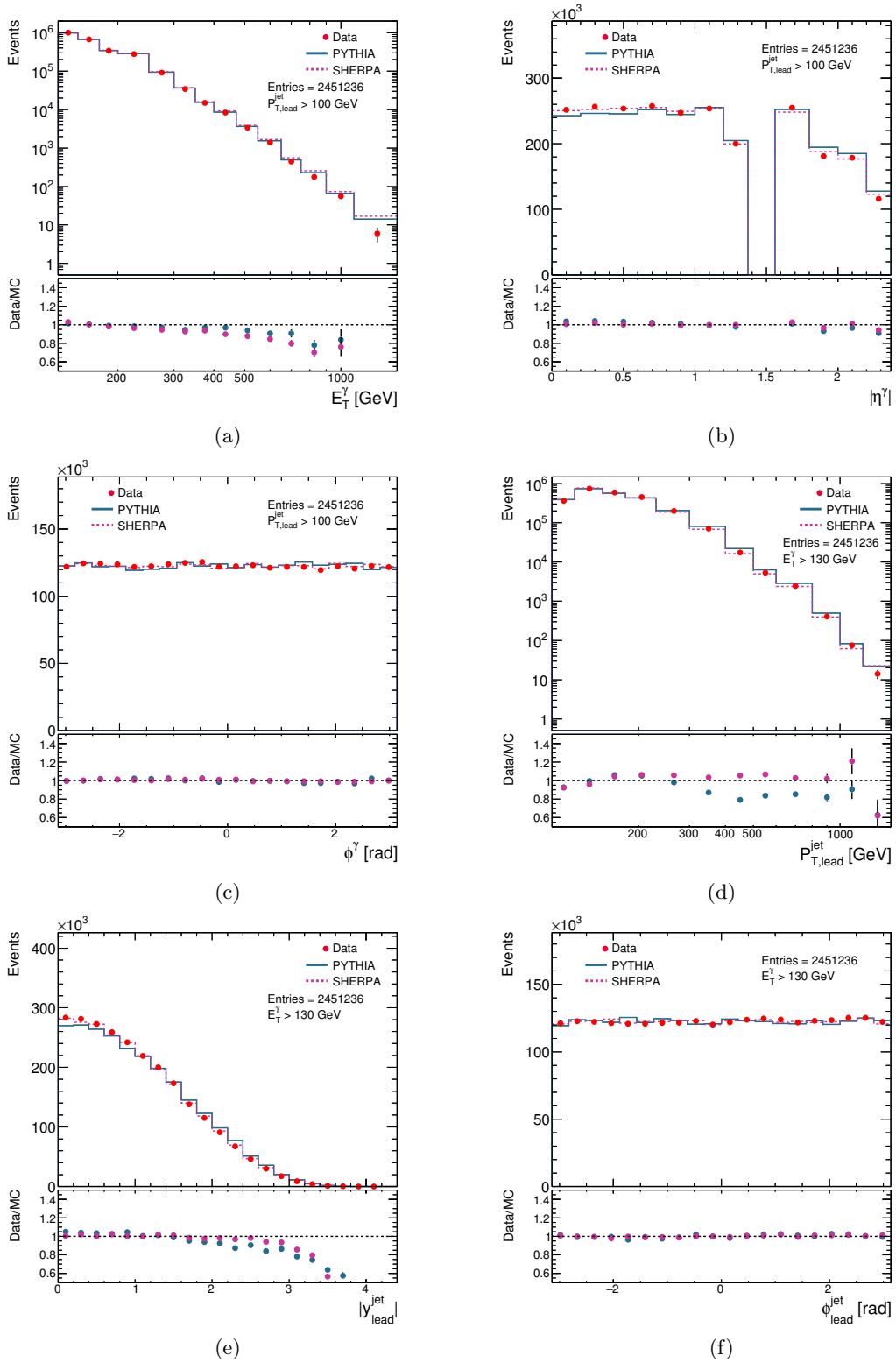
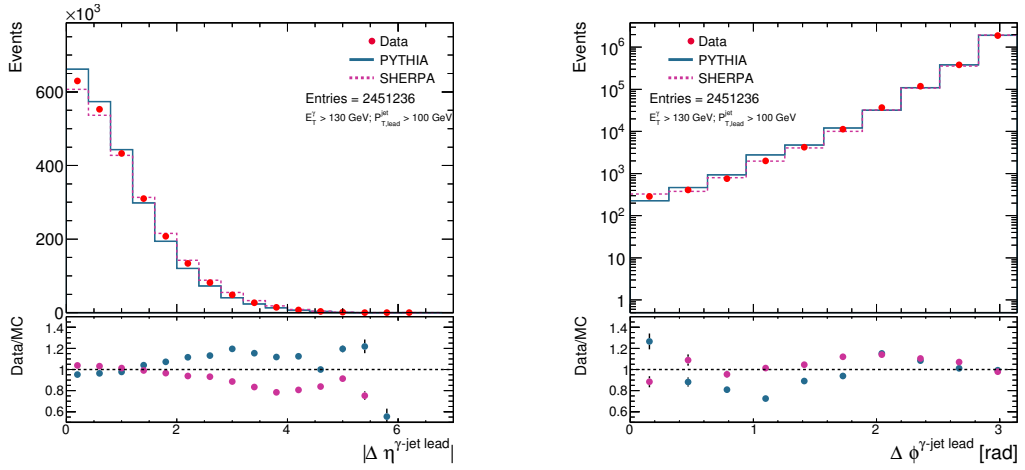
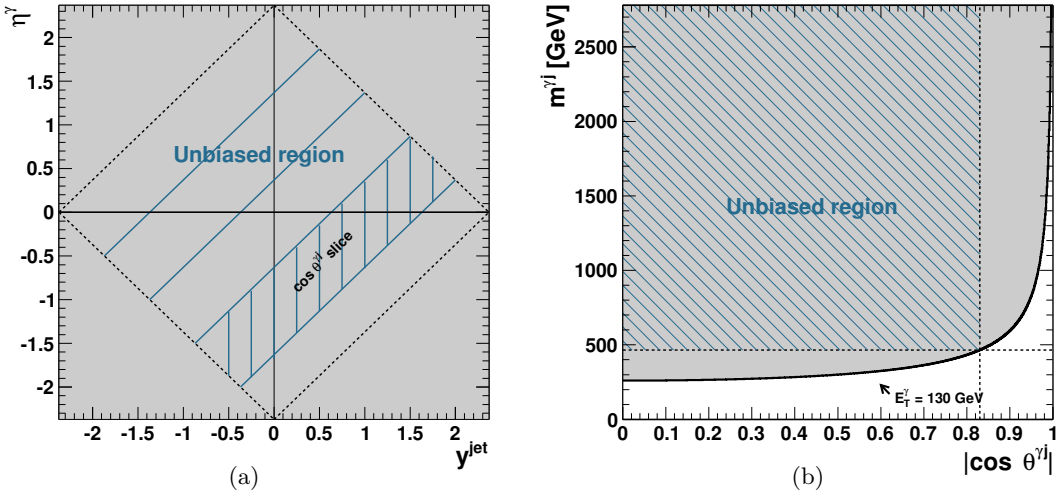


Figure 6.2: *The measured E^{γ}_{T} (a), $|\eta^{\gamma}|$ (b), ϕ^{γ} (c), $p^{\mathrm{jet}}_{\mathrm{T,lead}}$ (d), $|y^{\mathrm{jet}}_{\mathrm{lead}}|$ (e) and $\phi^{\mathrm{jet}}_{\mathrm{lead}}$ (f) distributions (dots). For comparison, the MC simulations of the signal from* PYTHIA *(blue histogram) and* SHERPA *(purple histogram) are also included. The MC distributions are normalised to the number of data events. The lower part of the figure shows the ratio of data and MC distributions.*



(a) (b)

Figure 6.3: *The measured* $\Delta\eta^{\gamma-\text{jet lead}}$ *(a) and* $\Delta\phi^{\gamma-\text{jet lead}}$ *(b) distributions. Details as in the caption to figure 6.2.*



(a) (b)

Figure 6.4: *The selected regions in the* η^{γ}*-*y^{jet} *(a) and* $m^{\gamma j}$*-*$|\cos\theta^{\gamma j}|$ *(b) planes. In (a), the dashed lines correspond to:* $\eta^{\gamma} + y^{\text{jet}} = 2.37$ *(first quadrant),* $\eta^{\gamma} - y^{\text{jet}} = 2.37$ *(second quadrant),* $\eta^{\gamma} + y^{\text{jet}} = -2.37$ *(third quadrant) and* $\eta^{\gamma} - y^{\text{jet}} = -2.37$ *(fourth quadrant). In (b), the horizontal (vertical) dashed line corresponds to* $m^{\gamma j} = 467$ *GeV* ($|\cos\theta^{\gamma j}| = 0.83$) *and the solid line corresponds to* $E_{\text{T}}^{\gamma} = 130$ *GeV.*

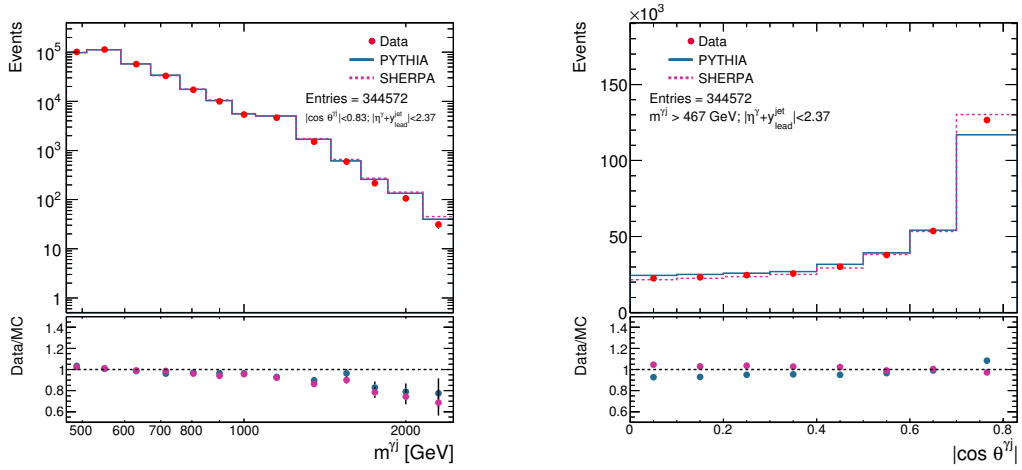
6.4.4 Colour coherence

In order to study colour coherence effects the observable β, which is defined in terms of the photon or leading jet was constructed. $\beta^{\text{jet}}_{\text{lead}}$ and β^{γ} is the angle of the sub-leading jet around the leading jet (photon) in the $\eta - \phi$ plane.

Colour coherence modifies how QCD radiation is distributed around a coloured parton (see section 1.5.4). By measuring the β observable for events in which the sub-leading jet is close to the leading jet or the photon, β provides a measurement of the direction of the QCD radiation around those objects with respect to the initial- or final-state partons.

The $\beta^{\text{jet}}_{\text{lead}}$ observable was reconstructed as

$$\beta^{\text{jet}}_{\text{lead}} = \tan^{-1}\left(|\Delta\phi^{\text{jet lead}-\text{jet sublead}}|, H^{\text{jet}}\right), \tag{6.1}$$



(a) (b)

Figure 6.5: *The measured $m^{\gamma j}$ (a) and $|\cos\theta^{\gamma j}|$ (b) distributions. Details as in the caption to figure 6.2.*

where $H^{\text{jet}} = \text{sign}(\eta^{\text{jet}}_{\text{lead}}) \cdot (\eta^{\text{jet}}_{\text{sublead}} - \eta^{\text{jet}}_{\text{lead}})$, and the β^{γ} observable was reconstructed as

$$\beta^{\gamma} = \tan^{-1}\left(|\Delta\phi^{\gamma,lead-jet\,sublead}|, H^{\gamma}\right), \tag{6.2}$$

where $H^{\gamma} = \text{sign}(\eta^{\gamma}) \cdot (\eta^{\text{jet}}_{\text{sublead}} - \eta^{\gamma})$. Using these definitions, 0 and π define the event plane between the leading jet or photon and the beam axis; $\beta = 0$ always points to the closest beam.

To construct the samples, events were required to have at least two jets with $p^{\text{jet}}_{\text{T}} > 50$ GeV and the pseudorapidity of the leading jet was restricted to $|\eta^{\text{jet}}_{\text{lead}}| < 2.37$ to make the comparison between β^{γ} and $\beta^{\text{jet}}_{\text{lead}}$ possible. In addition, only those events with $\Delta\phi^{\gamma-\text{jet lead}} > 2.6$ were kept. This last requirement aims to select back-to-back configurations.

$\beta^{\text{jet}}_{\text{lead}}$ was measured for events in which $1 < \Delta R^{\text{jet}}_{\text{lead jet}-\text{sublead jet}} < 1.5$ while β^{γ} was measured for events in which $1 < \Delta R^{\text{jet}}_{\gamma-\text{sublead jet}} < 1.5$. The lower limit was chosen because of the overlap requirement with the photon (no jet is allowed nearer than $\Delta R = 1$ to the photon, see section 6.3) while the upper limit was chosen to suppress events in which the rings around the leading jet and the photon overlap while maintaining high enough statistics.

There are 138578 events in the $\beta^{\text{jet}}_{\text{lead}}$ sample and 68446 events in the β^{γ} sample. Figure 6.9 shows schematic diagrams for the definitions of $\beta^{\text{jet}}_{\text{lead}}$ and β^{γ}. Figure 6.10 shows the $\beta^{\text{jet}}_{\text{lead}}$ and β^{γ} distributions.

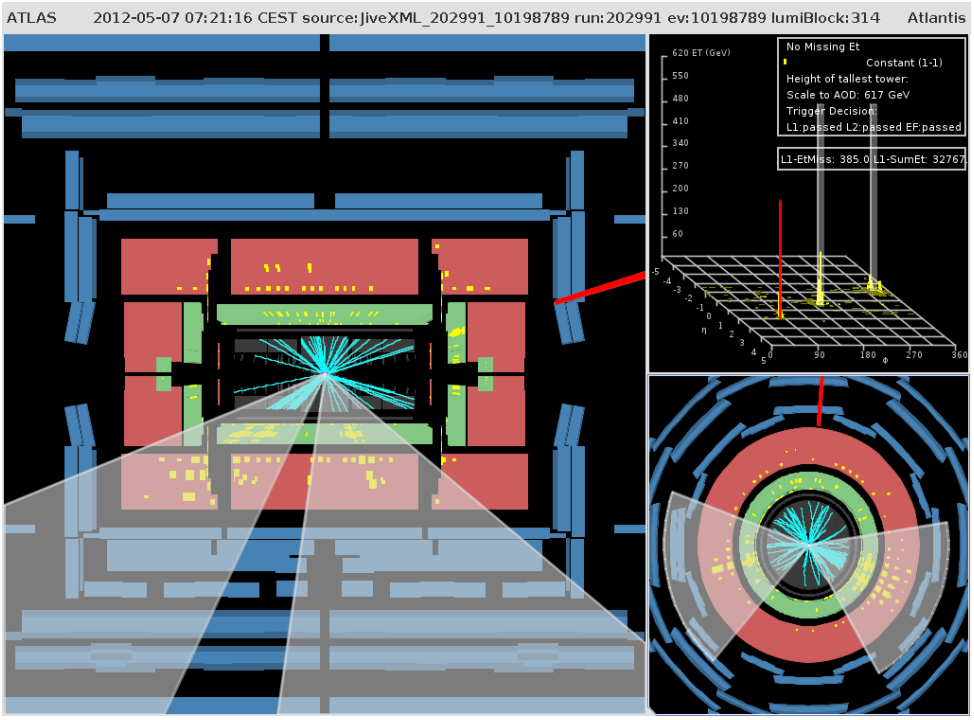


Figure 6.6: *Event display showing an event from the 2012 $\sqrt{s} = 8$ TeV dataset belonging to the photon+ two jets sample. Red lines represent isolated photons of tight quality, light blue lines represent tracks in the inner detector and yellow squares represent energy deposits in the calorimeters. Jets are represented as semi-transparent white regions. The EM calorimeter is drawn in green while the hadronic calorimeter is drawn in red.*

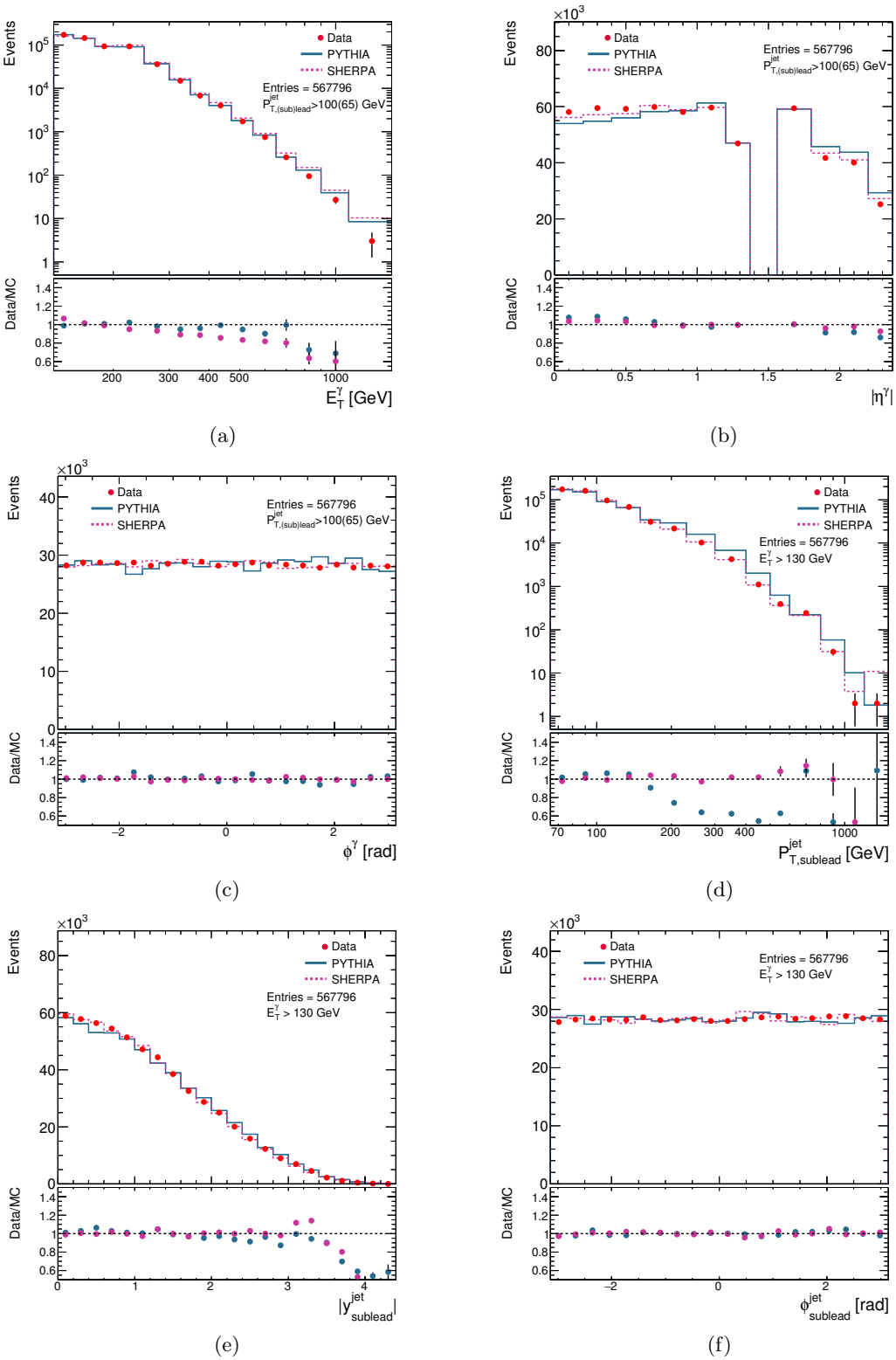


Figure 6.7: *The measured E^{γ}_{T} (a), $|\eta^{\gamma}|$ (b), ϕ^{γ} (c), $p^{\mathrm{jet}}_{\mathrm{T,sublead}}$ (d), $|y^{\mathrm{jet}}_{\mathrm{sublead}}|$ (e) and $\phi^{\mathrm{jet}}_{\mathrm{sublead}}$ (f) distributions for the photon+two-jets sample. Details as in the caption to figure 6.2.*

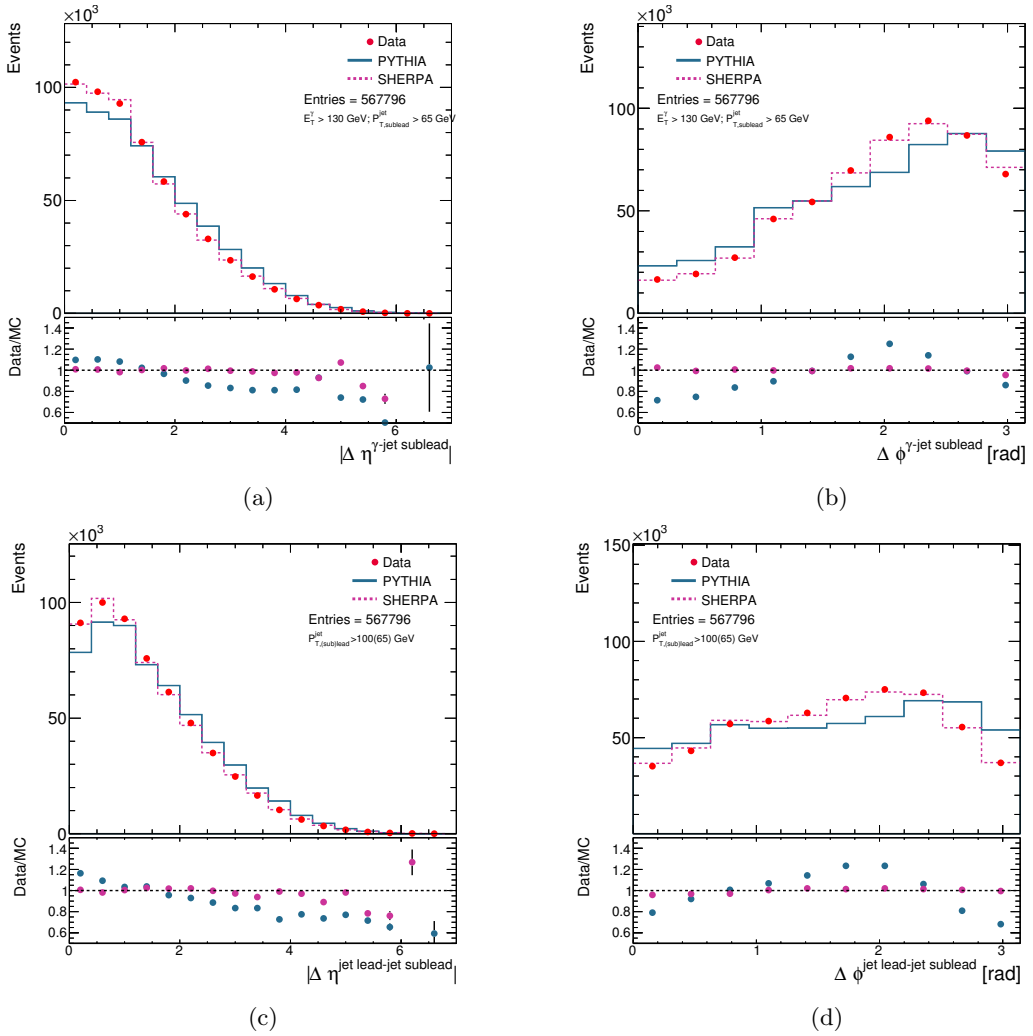
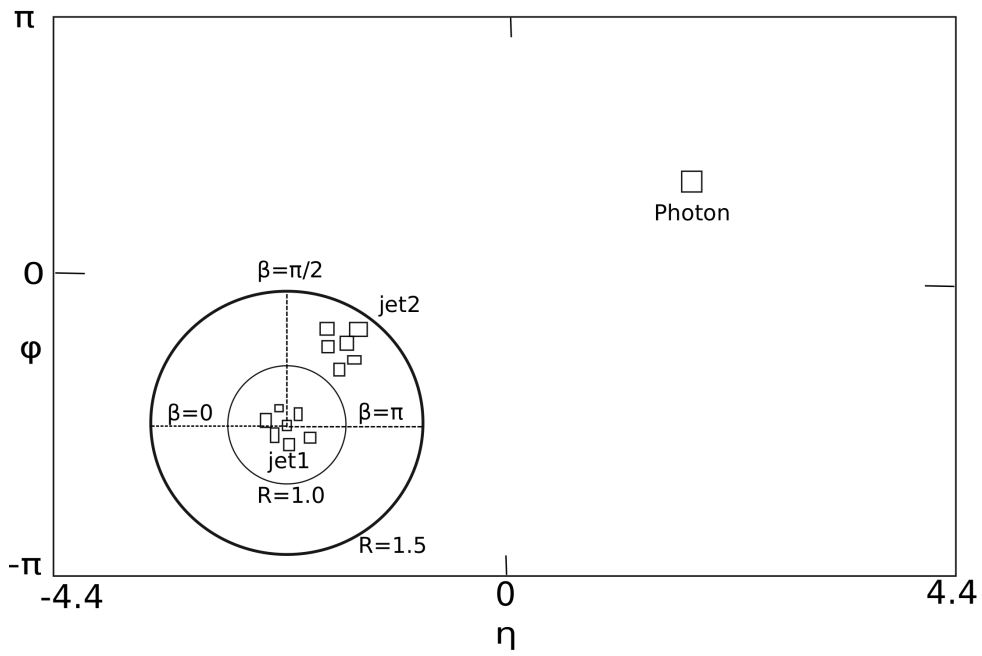
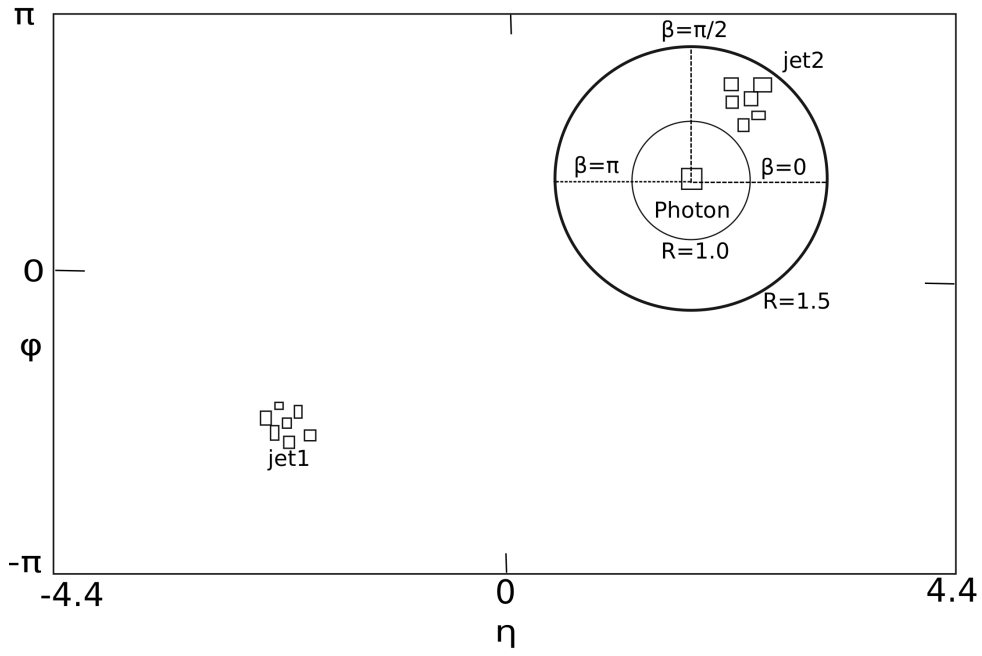


Figure 6.8: *The measured* $\Delta\eta^{\gamma-\text{jet sublead}}$ *(a),* $\Delta\phi^{\gamma-\text{jet sublead}}$ *(b),* $\Delta\eta^{\text{jet lead}-\text{jet sublead}}$ *(c) and* $\Delta\phi^{\text{jet lead}-\text{jet sublead}}$ *(d) distributions for the photon+two-jets sample. Details as in the caption to figure 6.2.*

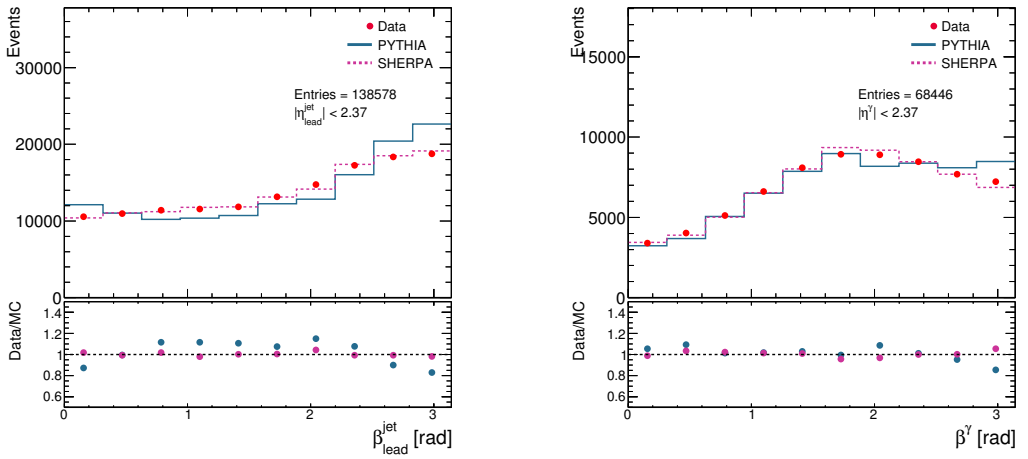


(a)



(b)

Figure 6.9: *Schematic diagrams that show the definitions of* $\beta^{\text{jet}}_{\text{lead}}$ *and* β^{γ}.



(a) (b)

Figure 6.10: *The measured $\beta^{\text{jet}}_{\text{lead}}$ (a) and β^{γ} (b) distributions. Details as in the caption to figure 6.2.*

6.4.5 photon+three-jets

The production of photon plus three additional jets was investigated by selecting events in which, in addition to the photon+two-jets requirements (see section 6.4.3), there is a third jet (sub-sub-leading jet) with $p^{\text{jet}}_{\text{T}} > 50$ GeV. The number of data events selected in this sample amounts to 164062. Figure 6.11 shows an event selected in this sample. The presence of a photon (red line) and three jets (semi-transparent white region) is seen.

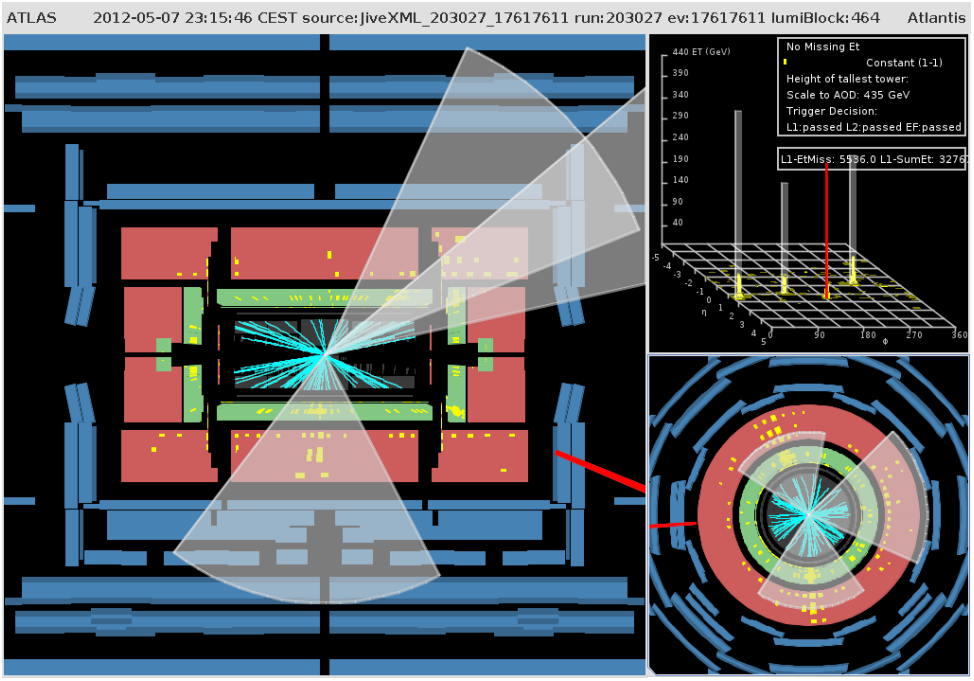


Figure 6.11: *Event display showing an event from the 2012 $\sqrt{s} = 8$ TeV dataset belonging to the photon+three jets sample. Red lines represent isolated photons of tight quality, light blue lines represent tracks in the inner detector and yellow squares represent energy deposits in the calorimeters. Jets are represented as semi-transparent white regions. The EM calorimeter is drawn in green while the hadronic calorimeter is drawn in red.*

Figure 6.12 show the E^{γ}_{T}, $|\eta^{\gamma}|$, ϕ^{γ}, the transverse-momentum of the sub-sub-leading jet

($p_{\mathrm{T,subsublead}}^{\mathrm{jet}}$), the rapidity of the sub-sub-leading jet ($|y_{\mathrm{subsublead}}^{\mathrm{jet}}|$) and the azimuthal angle of the sub-sub-leading jet ($\phi_{\mathrm{subsublead}}^{\mathrm{jet}}$) distributions. Figure 6.13 show the distributions for the pseudorapidity and azimuthal angle differences between the photon and the sub-sub-leading jet ($\Delta\eta^{\gamma-\mathrm{jet\,subsublead}}$ and $\Delta\phi^{\gamma-\mathrm{jet\,subsublead}}$), between the leading jet and the sub-sub-leading jet ($\Delta\eta^{\mathrm{jet\,lead}-\mathrm{jet\,subsublead}}$ and $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,subsublead}}$) and between the sub-leading jet and the sub-sub-leading jet ($\Delta\eta^{\mathrm{jet\,sublead}-\mathrm{jet\,subsublead}}$ and $\Delta\phi^{\mathrm{jet\,sublead}-\mathrm{jet\,subsublead}}$).

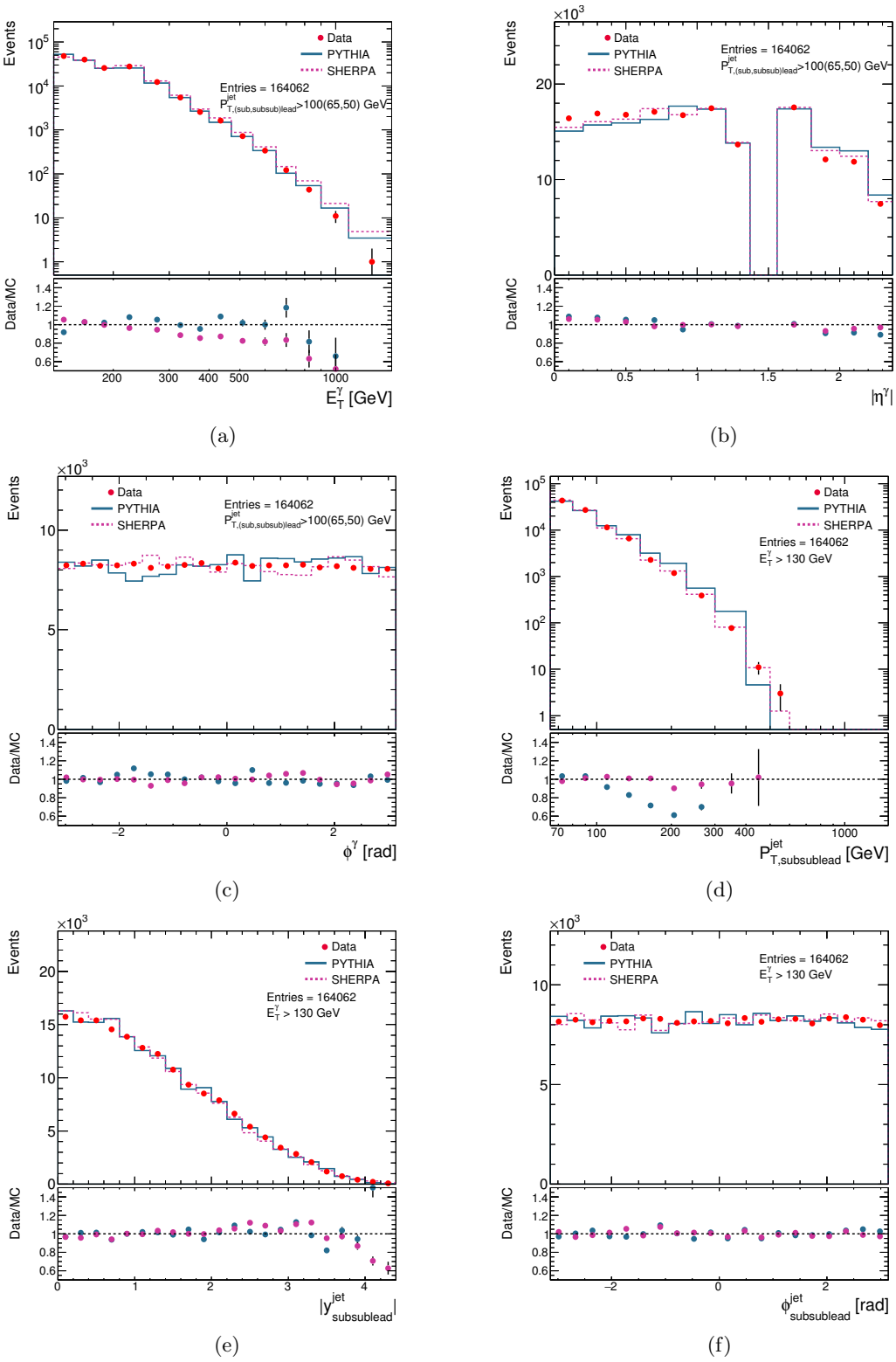


Figure 6.12: *The measured E^{γ}_{T} (a), $|\eta^{\gamma}|$ (b), ϕ^{γ} (c), $p^{\mathrm{jet}}_{\mathrm{T,subsublead}}$ (d), $|y^{\mathrm{jet}}_{\mathrm{subsublead}}|$ (e) and $\phi^{\mathrm{jet}}_{\mathrm{subsublead}}$ (f) distributions for the photon+three-jets sample. Details as in the caption to figure 6.2.*

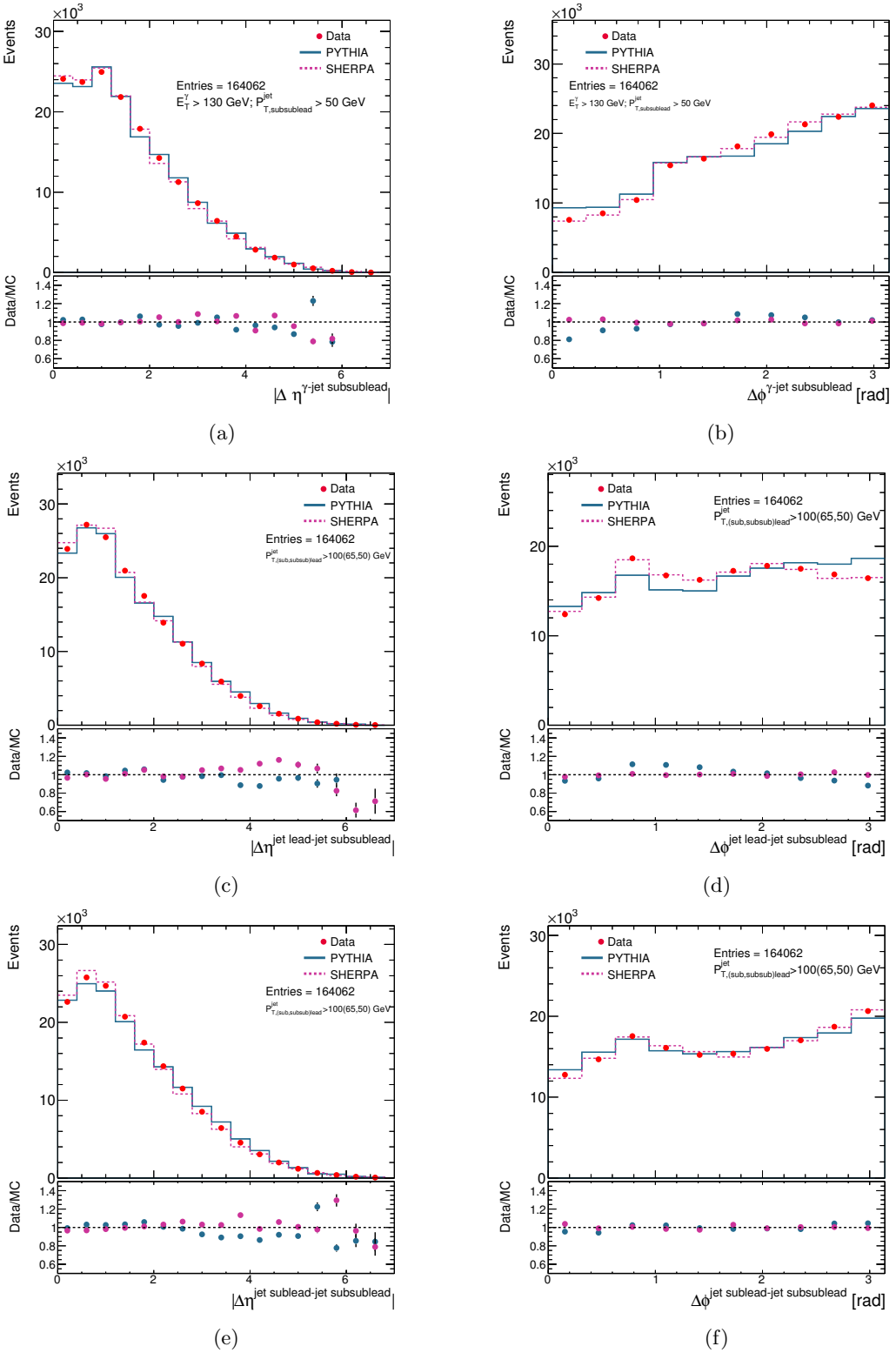


Figure 6.13: *The measured* $\Delta\eta^{\gamma-\text{jet subsublead}}$ *(a),* $\Delta\phi^{\gamma-\text{jet subsublead}}$ *(b),* $\Delta\eta^{\text{jet lead}-\text{jet subsublead}}$ *(c),* $\Delta\phi^{\text{jet lead}-\text{jet subsublead}}$ *(d),* $\Delta\eta^{\text{jet sublead}-\text{jet subsublead}}$ *(e) and* $\Delta\phi^{\text{jet sublead}-\text{jet subsublead}}$ *(f) distributions for the photon+three-jets sample. Details as in the caption to figure 6.2.*

6.5 Additional corrections to MC events

Additional corrections were applied to the simulated events to match the overall event conditions of the data sample and to account for known discrepancies between data and simulation. These corrections are:

- pile-up corrections: to match the in-time and out-of-time pile-up conditions in the data, the distribution of $\langle\mu\rangle$ in simulated events was reweighted to that of the data, after applying a general scale factor of 1.09 to achieve a better agreement between data and MC for the distribution on the number of primary vertices. Figure 6.14 shows the distribution of $\langle\mu\rangle$ for data and PYTHIA and SHERPA MC before and after this reweighting was applied. The reweighting factors for both MC simulations are very similar;

- corrections to the photon identification:

  - fudge factors were applied to the photon discriminating variables in the signal MC simulated events to correct the efficiency for the differences observed between data and MC simulation in the photon-cluster discriminating variables;
  - scale factors were applied to tight photons in MC simulated events to account for differences between data and MC simulation;

- corrections to $E_{\mathrm{T}}^{\mathrm{iso}}$: differences on the modelling of the signal isolation distribution are observed between data and MC simulations (see figure 6.15 and 6.16). To take into account these differences, the $E_{\mathrm{T}}^{\mathrm{iso}}$ variable in the MC events was corrected event-by-event by a value $\Delta(\eta^{\gamma}, E_{\mathrm{T}}^{\gamma})$ so that the mean value of the distribution coincides with that of the data. The correction was computed as $\Delta(\eta^{\gamma}, E_{\mathrm{T}}^{\gamma}) = p_0^{\mathrm{data}} - p_0^{\mathrm{MC}}$, where p_0^{data} and p_0^{MC} are the p_0 values obtained from a fit to the $E_{\mathrm{T}}^{\mathrm{iso}}$ distribution in data and MC, respectively, using the functional form:

$$\begin{aligned} f(x, p_0, p_1, p_2, p_3) &= \frac{N_2}{N_1}\, e^{-0.5((x-p_0)/p_1)^2} \text{ if } x < p_3, \\ &= e^{-(p_2+bx)} \text{ if } x \geq p_3, \end{aligned}$$

  where $b = (p_3 - p_0)/p_1^2$, $N_1 = e^{-0.5((p_3-p_0)/p_1)^2}$ and $N_2 = e^{-(p_2+bp_3)}$ to ensure the continuity and derivability of the function in $x = p_3$. The data distribution was obtained after background subtraction (see Section 6.6). Figures 6.17 and 6.18 show the improvement in the agreement between data and MC simulations after this correction was applied;

- correction to the Z-vertex distribution. The Z-vertex distribution in the MC samples does not describe the data well. A reweigthing procedure as a function of the Z-vertex position at particle level was implemented, separately for PYTHIA and SHERPA. Figure 6.19 shows the Z-vertex distribution for data and MC simulations before and after the reweigthing. A good description of the data is obtained after the reweigthing.

Cross-checks of the effects of all these additional corrections to the simulated events on the measured cross sections are included in Section 7.6.

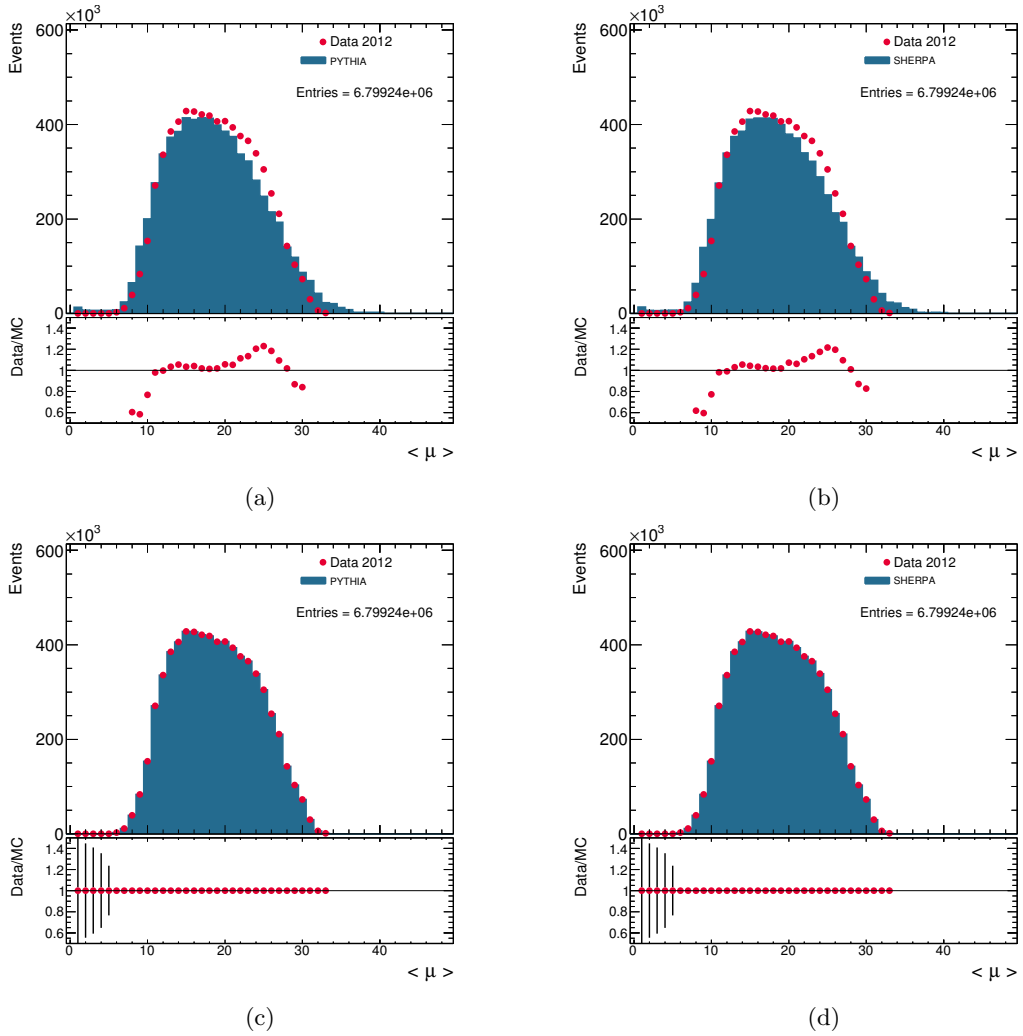


Figure 6.14: *Distribution of the average number of interactions per bunch crossing in data (dots) and* Pythia *and* Sherpa *MC (histograms) before (a,b) and after (c,d) reweighting. The lower part of the plots shows the ratio between data and MC.*

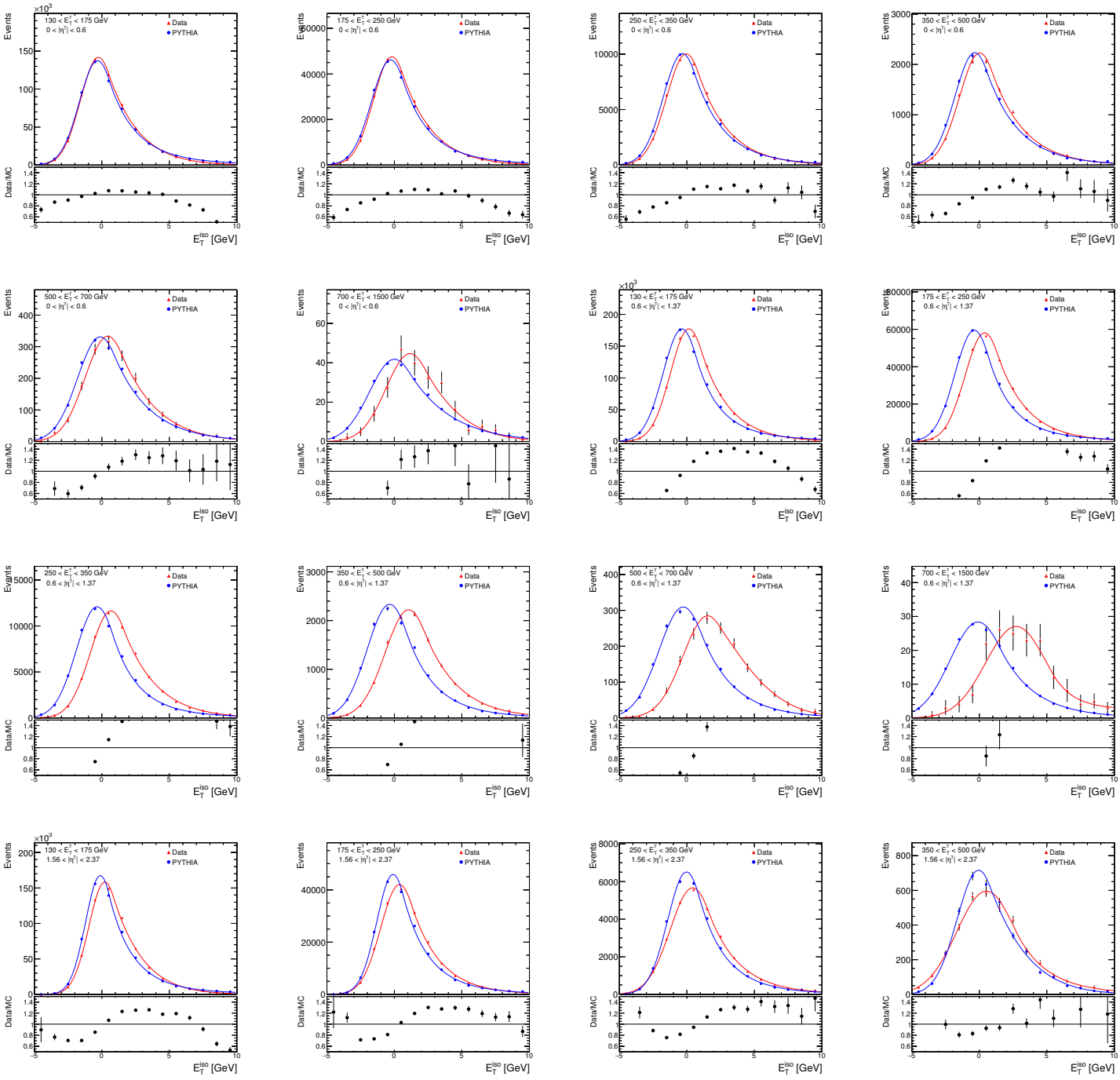


Figure 6.15: *$E_{\mathrm{T}}^{\mathrm{iso}}$ distribution after background subtraction (see text) for photon candidates in data (red dots) in different E_{T}^{γ} and $|\eta^{\gamma}|$ regions. For comparison, the MC simulations of the signal from* PYTHIA *(blue dots) are also included. The MC distributions are normalised to the number of data events in each region. The curves are the fits performed to data and MC events, using the function as described in the text. The lower part of the figure shows the ratio of data and MC distributions.*

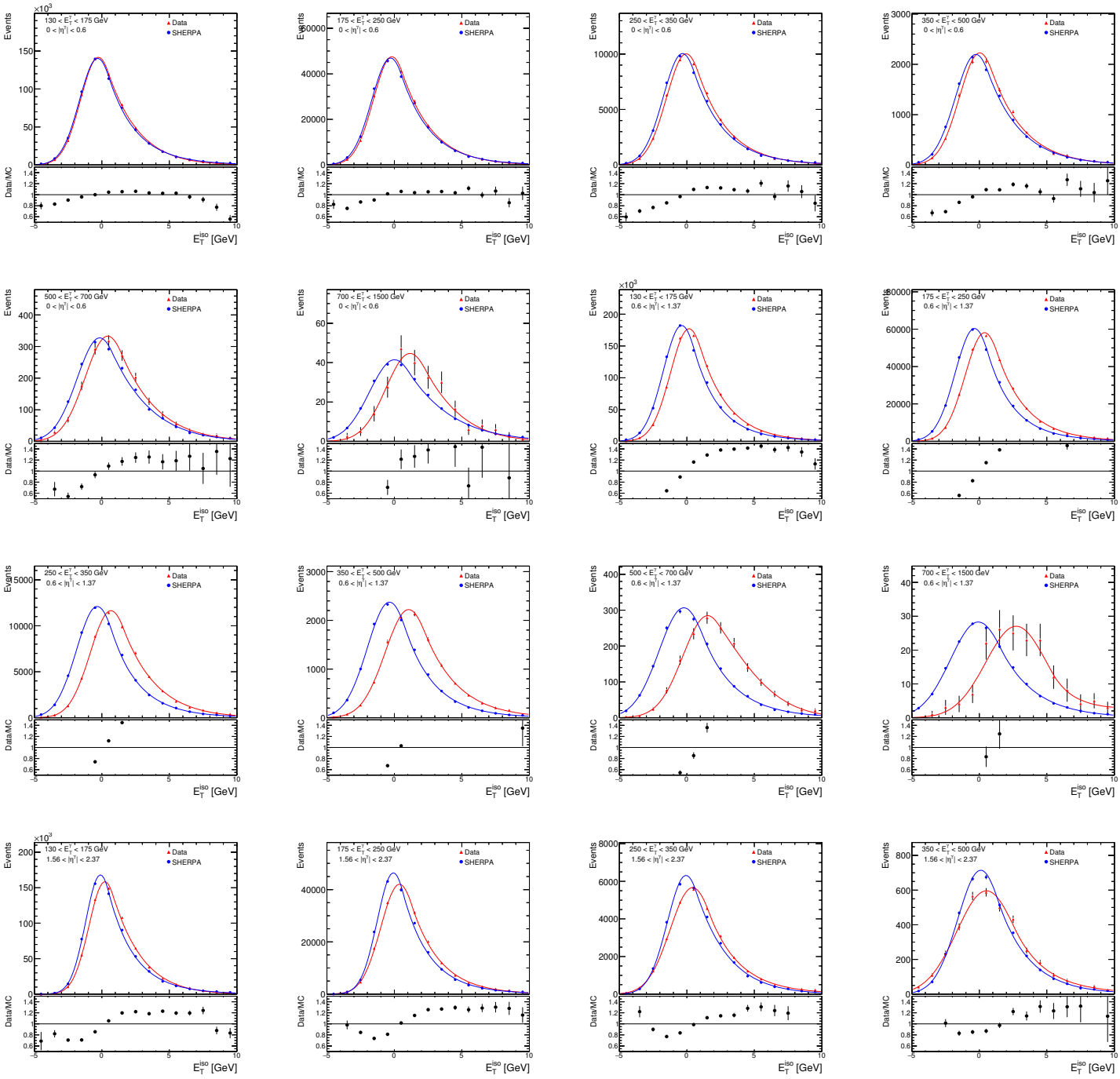


Figure 6.16: $E_{\mathrm{T}}^{\mathrm{iso}}$ *distribution after background subtraction (see text) for photon candidates in data (red dots) in different* E_{T}^{γ} *and* $|\eta^{\gamma}|$ *regions. For comparison, the MC simulations of the signal from* SHERPA *(blue dots) are also included. Other details as in the caption to figure 6.15.*

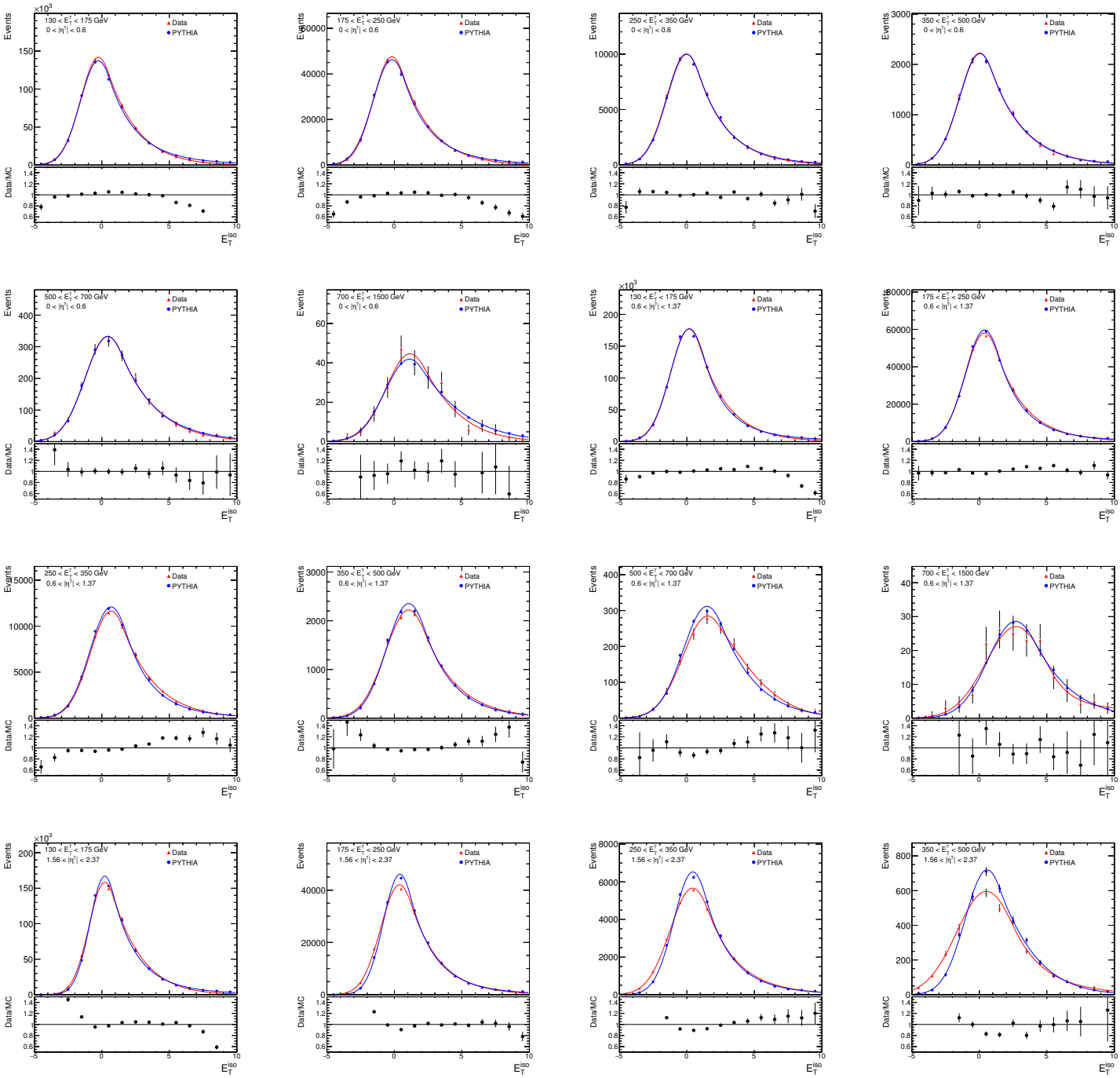


Figure 6.17: $E_{\mathrm{T}}^{\mathrm{iso}}$ *distribution after background subtraction (see text) for photon candidates in data (red dots) in different* E_{T}^{γ} *and* $|\eta^{\gamma}|$ *regions. For comparison, the MC simulations of the signal from* PYTHIA *(blue dots) after the correction described in the text are also included. Other details as in the caption to figure 6.15.*

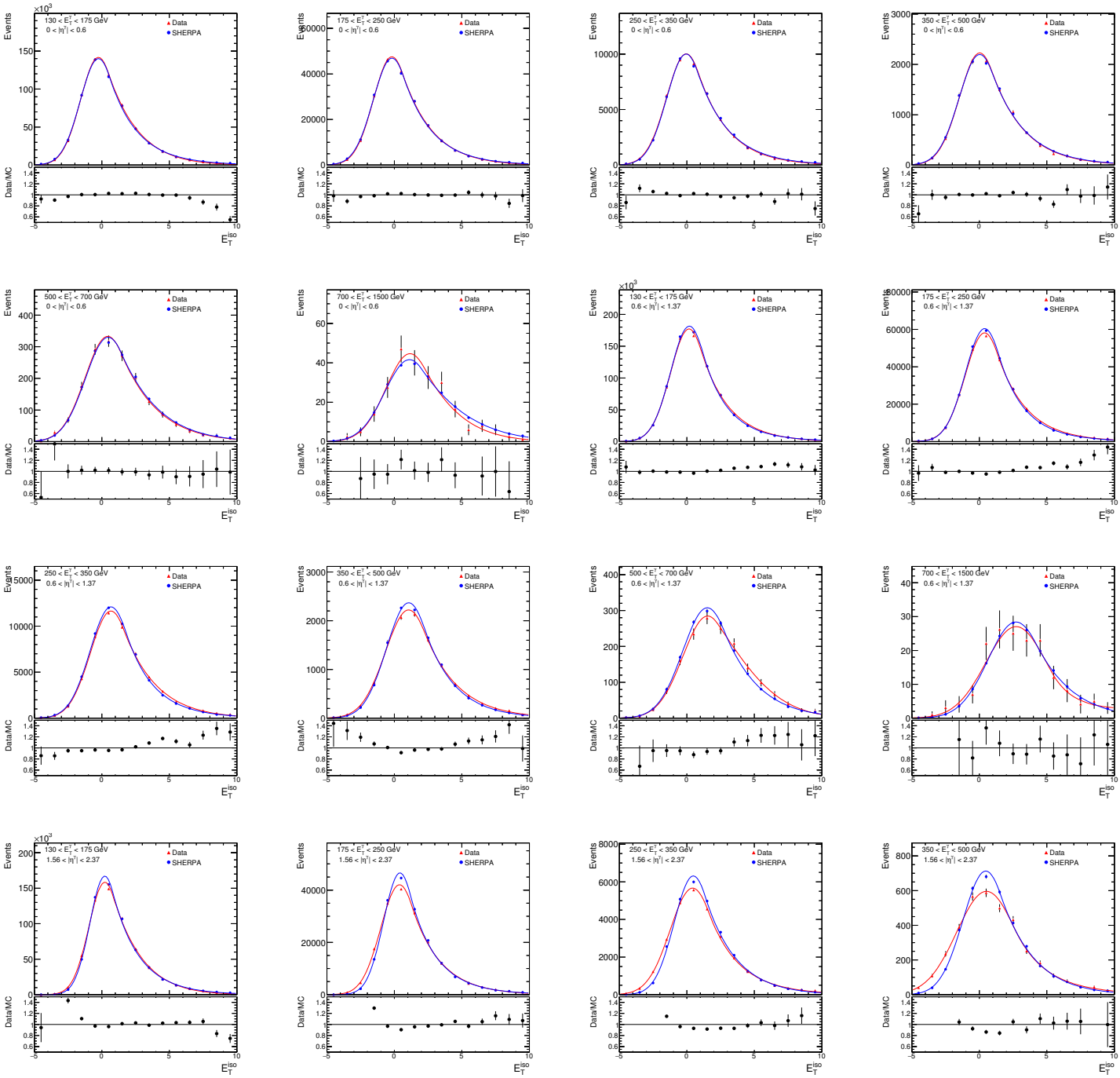


Figure 6.18: $E_{\mathrm{T}}^{\mathrm{iso}}$ *distribution after background subtraction (see text) for photon candidates in data (red dots) in different* E_{T}^{γ} *and* $|\eta^{\gamma}|$ *regions. For comparison, the MC simulations of the signal from* SHERPA *(blue dots) after the correction described in the text are also included. Other details as in the caption to figure 6.15.*

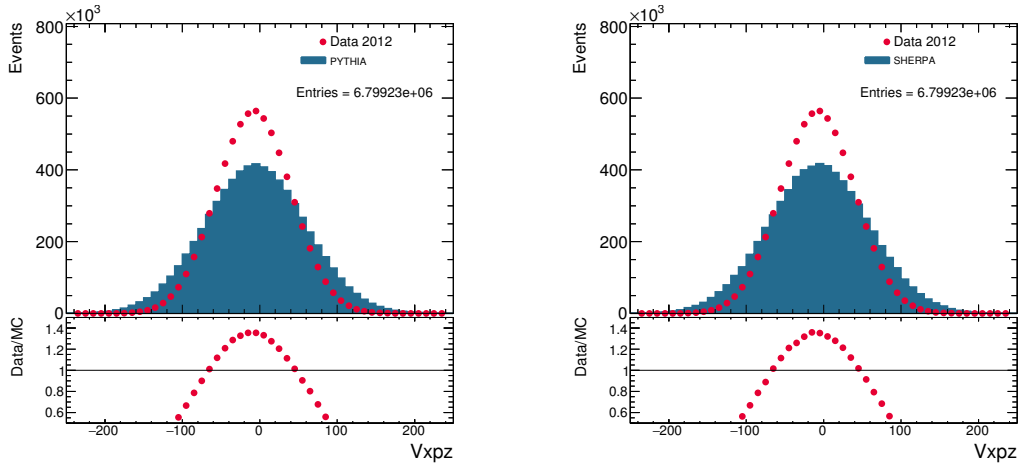


Figure 6.19: *Z-vertex distribution for data (dots) and* Pythia *and* Sherpa *MC (histograms) before (a,b) and after (c,d) reweighting. The lower part of the plots shows the ratio between data and MC.*

6.6 Background estimation and subtraction

A non-negligible contribution of background remains in the selected sample, even after the application of the tight identification and isolation requirements. This background comes predominantly from QCD processes, in which a jet is misidentified as a photon. This jet contains usually a light neutral meson, predominantly a π^0 that decays into two collimated photons, which carries most of the energy of the jet.

The isolation profile of a signal-enriched sample was extracted from the data. For this study, a sample was obtained by applying all the selection criteria described in Section 6.3, except for the tight identification and isolation requirements. Two subsamples were selected by applying either the tight or the non-tight identification criteria; the subsample with non-tight identification criteria is expected to be enriched in background candidates. The $E_{\mathrm{T}}^{\mathrm{iso}}$ distributions for these subsamples are shown in figure 6.20a. The $E_{\mathrm{T}}^{\mathrm{iso}}$ distribution of the non-tight candidates was scaled so that the integral for $E_{\mathrm{T}}^{\mathrm{iso}} > 10$ GeV, where the contribution from the signal is expected to be negligible, matched that of the tight candidates. The rescaled background distribution was subtracted from that of the tight photon candidates to extract the isolation profile of signal-like candidates, as shown in figure 6.20b; this isolation profile is compared to the MC simulations of Pythia and Sherpa in figure 6.20. A good description of the data by the MC simulations is obtained after this data-only background subtraction.

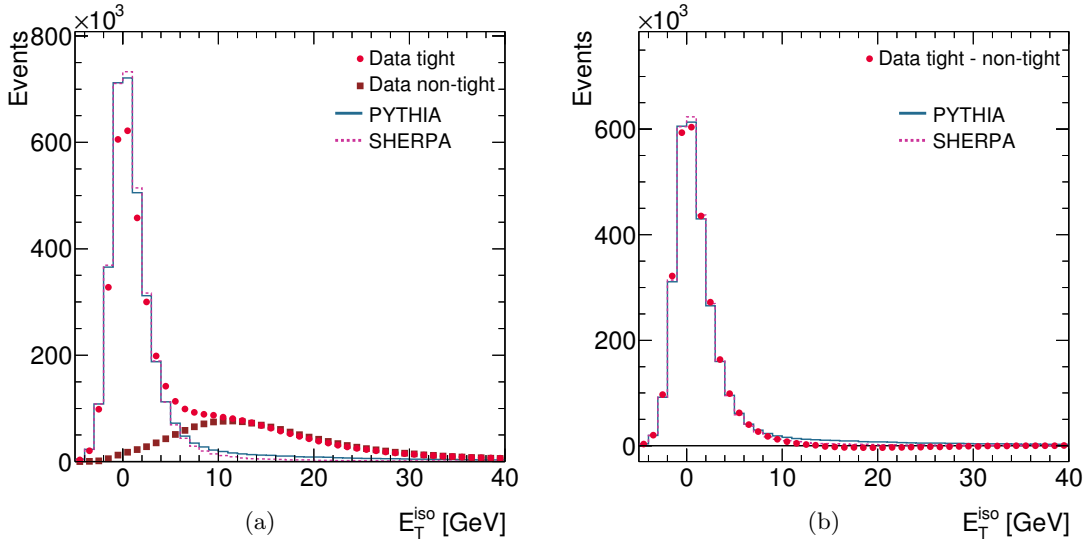


Figure 6.20: *(a) The measured $E_{\mathrm{T}}^{\mathrm{iso}}$ distribution before the isolation requirement and after applying the tight identification requirement (dots) and for those events which fail the tight identification ("non-tight") (squares). The "non-tight" distribution was normalised so that the integral of the "tight" and "non-tight" distributions for $E_{\mathrm{T}}^{\mathrm{iso}} > 10$ GeV coincides. For comparison, the MC simulations of the signal from* Pythia *(blue histogram) and* Sherpa *(purple histogram) are also included. The MC distributions are normalised to the number of data events in the tight distribution for $E_{\mathrm{T}}^{\mathrm{iso}} < 10$ GeV. (b) The measured $E_{\mathrm{T}}^{\mathrm{iso}}$ distribution before the isolation requirement and after applying the tight identification requirement and after subtracting the non-tight events (dots).*

The estimation of the background from the MC samples is not reliable. Therefore, a background-subtraction method, which does not rely on MC samples of background, was devised which uses signal-suppressed control regions to obtain a purer photon signal. The "2D-sideband" subtraction method [89] was used. The method is briefly explained below.

The background contamination was estimated and then subtracted by using a counting technique based on the observed number of events in control regions of a two-dimensional plane. This plane was defined by using the photon identification variable (γ_{ID}) and the $E_{\mathrm{T}}^{\mathrm{iso}}$ variable, as shown in figure 6.21. Four regions were defined in this plane:

- "A" is the signal region, which contains tight and isolated ($E_{\mathrm{T}}^{\mathrm{iso}} < 4.8 + 4.2 \cdot 10^{-3} \times E_{\mathrm{T}}^{\gamma}$ GeV) photon candidates;
- "B" is the control region with non-isolated background, which contains tight and non-isolated ($E_{\mathrm{T}}^{\mathrm{iso}} > 4.8 + 4.2 \cdot 10^{-3} \times E_{\mathrm{T}}^{\gamma} + 2$ GeV) photon candidates;
- "C" is the control region with non-identified background, which contains isolated ($E_{\mathrm{T}}^{\mathrm{iso}} < 4.8 + 4.2 \cdot 10^{-3} \times E_{\mathrm{T}}^{\gamma}$ GeV) and non-tight photon candidates;
- "D" is the background control region, which contains non-isolated ($E_{\mathrm{T}}^{\mathrm{iso}} > 4.8 + 4.2 \cdot 10^{-3} \times E_{\mathrm{T}}^{\gamma} + 2$ GeV) and non-tight photon candidates.

Therefore, the number of signal events in the signal region A is

$$N_A^{\mathrm{sig}} = N_A - R^{\mathrm{bg}} \cdot (N_B - \epsilon_B N_A^{\mathrm{sig}}) \cdot \frac{(N_C - \epsilon_C N_A^{\mathrm{sig}})}{(N_D - \epsilon_D N_A^{\mathrm{sig}})}, \tag{6.3}$$

where N_A^{sig} is the expected number of signal events, N_K with $K = A, B, C, D$ is the number of observed events in each region and

$$R^{\mathrm{bg}} = \frac{N_A^{\mathrm{bg}} \cdot N_D^{\mathrm{bg}}}{N_B^{\mathrm{bg}} \cdot N_C^{\mathrm{bg}}}$$

was taken as $R^{\mathrm{bg}} = 1$ for the nominal results; N_K^{bg} with $K = A, B, C, D$ is the number of background events in each region. Deviations with respect to $R^{\mathrm{bg}} = 1$ were taken as systematic uncertainties (see section 6.11). Equation 6.3 takes into account the expected number of signal events in the three background control regions via the signal leakage fractions, $\epsilon_K = N_K^{\mathrm{sig}}/N_A^{\mathrm{sig}}$ with $K = B, C, D$. The signal leakage fractions were extracted from the MC simulations of the signal and are shown in figures 6.22 to 6.29 for Pythia and Sherpa for the observables studied. The frations ϵ_c, which represent the signal leaking into the non-tight and isolated control region, are approximately constant as functions of all observables and around 2×10^{-2} and very similar for Pythia and Sherpa. The fractions ϵ_B, signal leakage into the tight and non-isolated control region, have similar shapes for Pythia and Sherpa as functions of all observables and their values are somewhat smaller for Sherpa; this is due to the different treatment of the brem component in Pythia and Sherpa. The fractions ϵ_D, the signal leakage into the non-tight and non-isolated control region, are very different for Pythia and Sherpa; this is the control region most affected by the different treatment of the brem component in both models.

The signal yield was determined from the observed yields in the data in the four regions of the γ_{ID} vs. $E_{\mathrm{T}}^{\mathrm{iso}}$ plane and the signal leakage fractions determined from the simulated events using equation 6.3. The signal purity, computed as $P = N_A^{\mathrm{sig}}/N_A$, is shown in figures 6.22 to 6.29 for Pythia and Sherpa MC for the observables studied. The purity is above 90% and using either Pythia or Sherpa to compute the signal leakage fractions yield very similar results.

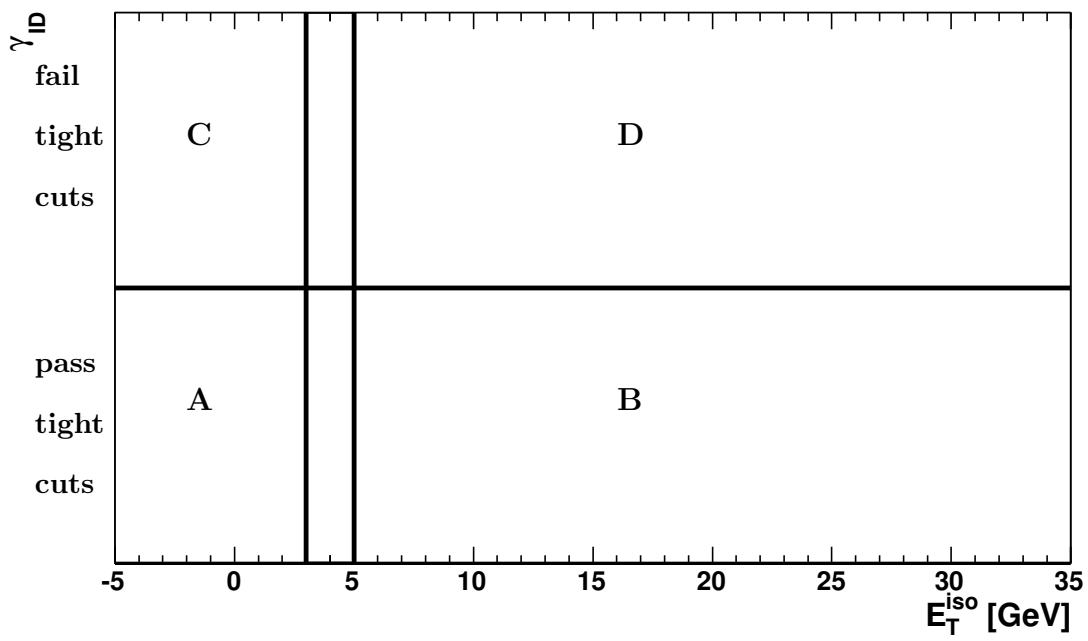


Figure 6.21: *Illustration of the two-dimensional plane of the photon identification variables vs. the transverse isolation energy used to estimate the background yield in the signal region, A, from the observed yields in the three control regions, B, C and D.*

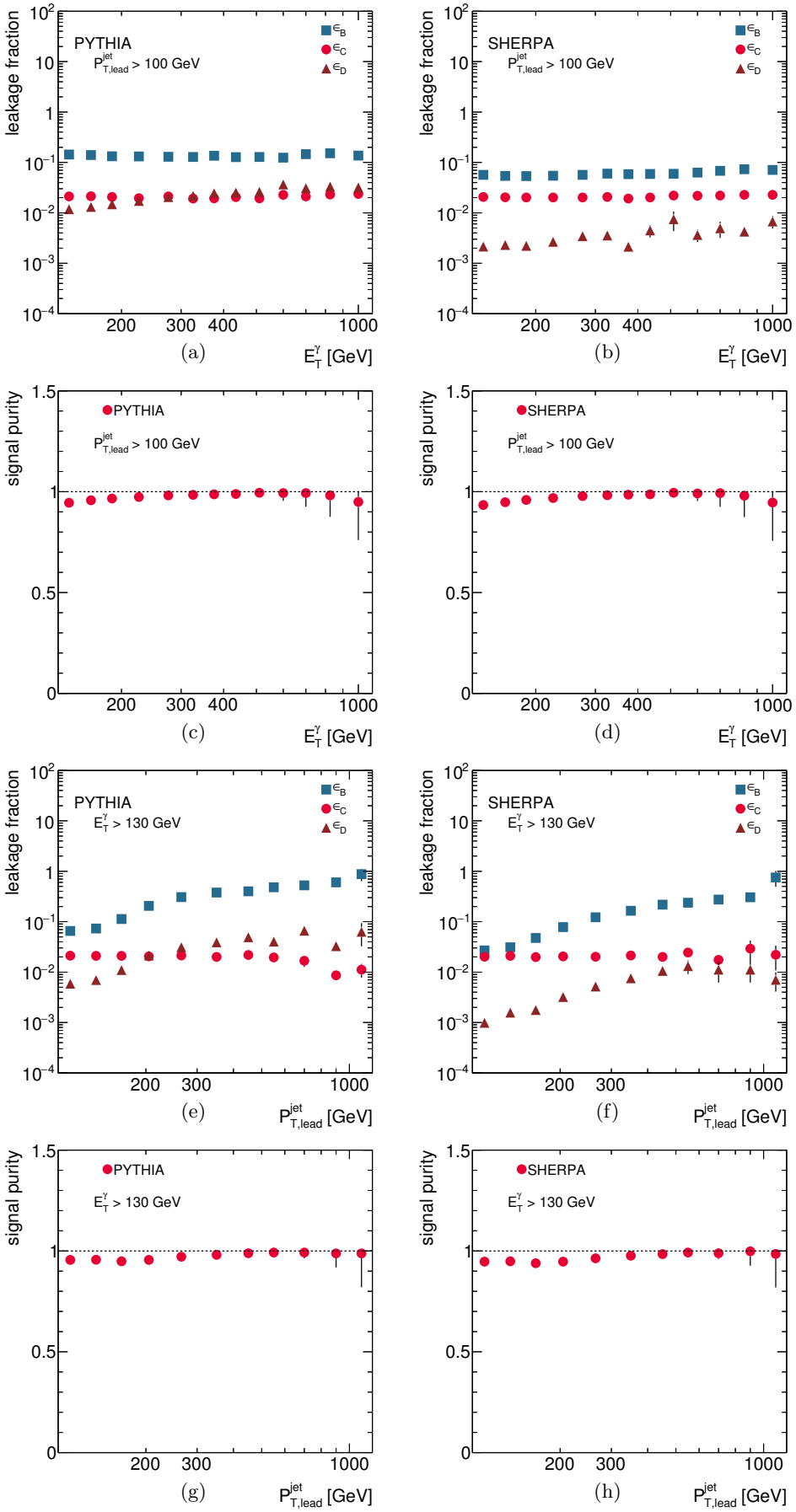


Figure 6.22: *Signal leakage fractions from* PYTHIA *(a,e) and* SHERPA *(b,f) for the B (squares), C (dots) and D (triangles) control regions as functions of* E_{T}^{γ} *(a,b) and* $p_{\mathrm{T,lead}}^{\mathrm{jet}}$ *(e,f) for the photon+one-jet sample. Estimated signal purities in data using signal leakage fractions from* PYTHIA *(c,g) and* SHERPA *(d,h) as functions of* E_{T}^{γ} *(c,d) and* $p_{\mathrm{T,lead}}^{\mathrm{jet}}$ *(g,h) for the photon+one-jet sample. In all plots, the error bars are statistical only.*

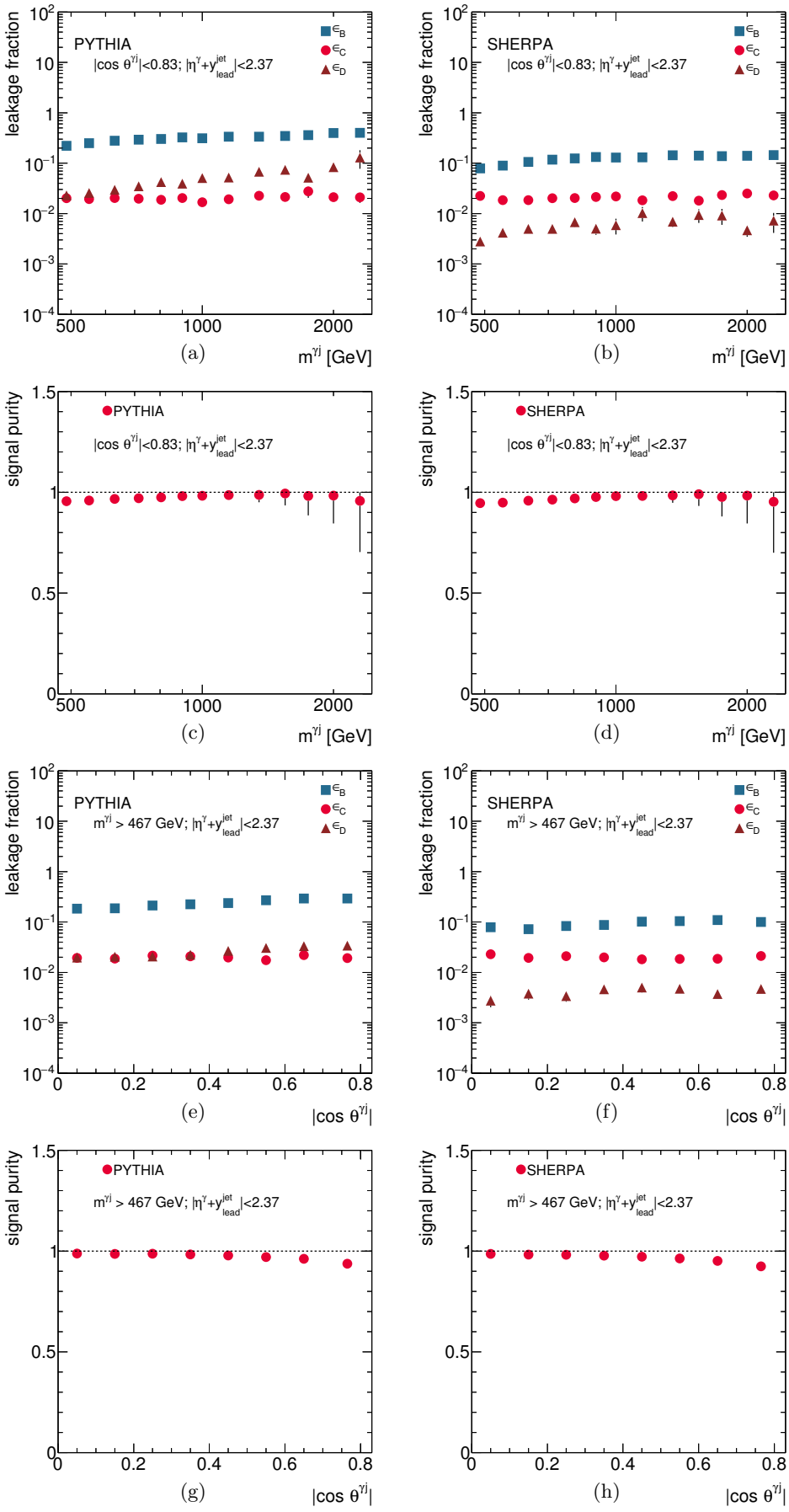


Figure 6.23: *Signal leakage fractions from* PYTHIA *(a,e) and* SHERPA *(b,f) for the B (squares), C (dots) and D (triangles) control regions as functions of* $m^{\gamma j}$ *(a,b) and* $\cos\theta^{\gamma j}$ *(e,f) for the photon+one-jet sample. Estimated signal purities in data using signal leakage fractions from* PYTHIA *(c,g) and* SHERPA *(d,h) as functions of* $m^{\gamma j}$ *(c,d) and* $\cos\theta^{\gamma j}$ *(g,h) for the photon+one-jet sample. In all plots, the error bars are statistical only.*

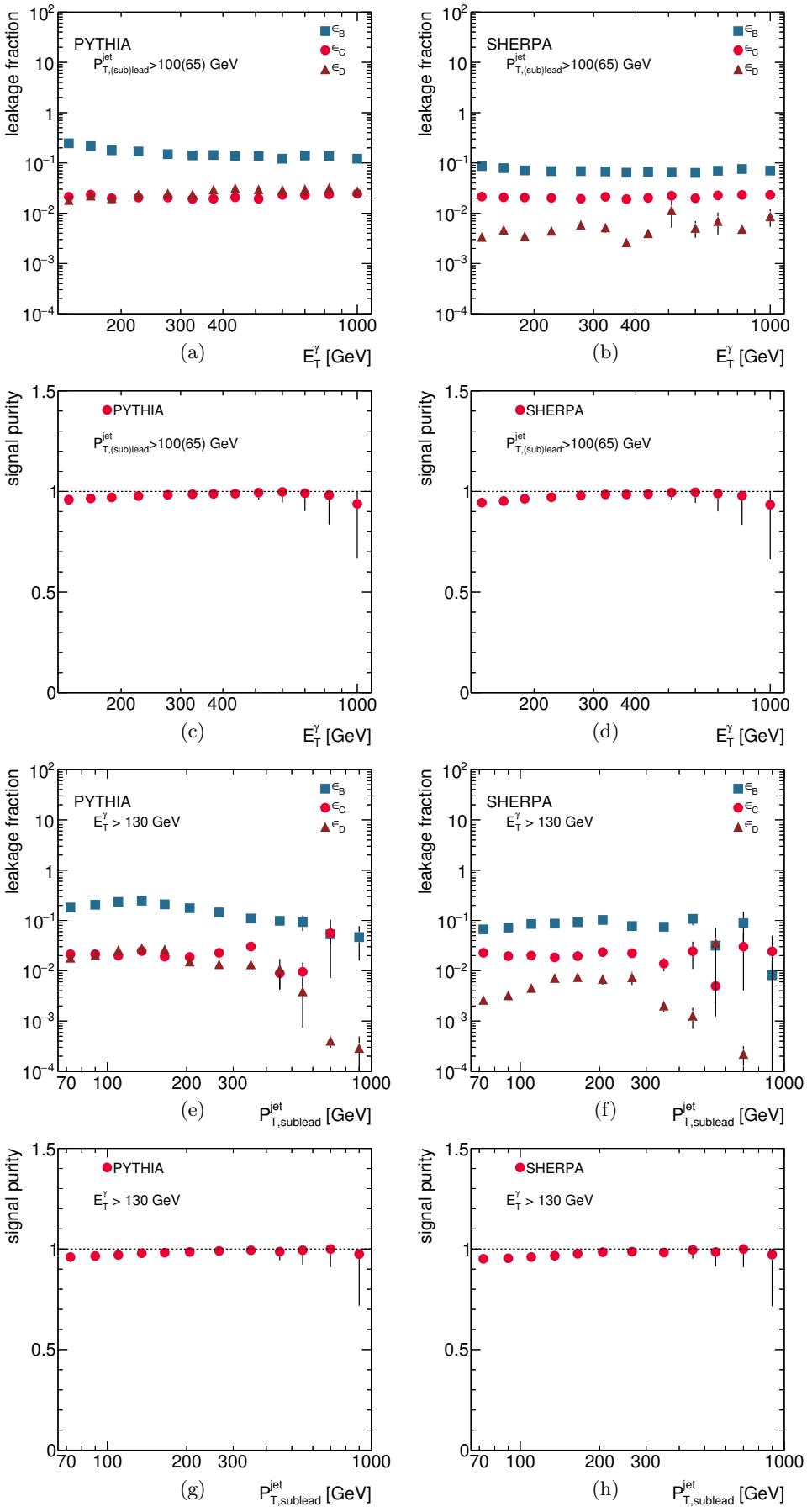


Figure 6.24: *Signal leakage fractions from* PYTHIA *(a,e) and* SHERPA *(b,f) for the B (squares), C (dots) and D (triangles) control regions as functions of* E^{γ}_{T} *(a,b) and* $p^{\mathrm{jet}}_{\mathrm{T,sublead}}$ *(e,f) for the photon+two-jets sample. Estimated signal purities in data using signal leakage fractions from* PYTHIA *(c,g) and* SHERPA *(d,h) as functions of* E^{γ}_{T} *(c,d) and* $p^{\mathrm{jet}}_{\mathrm{T,sublead}}$ *(g,h) for the photon+two-jets sample. In all plots, the error bars are statistical only.*

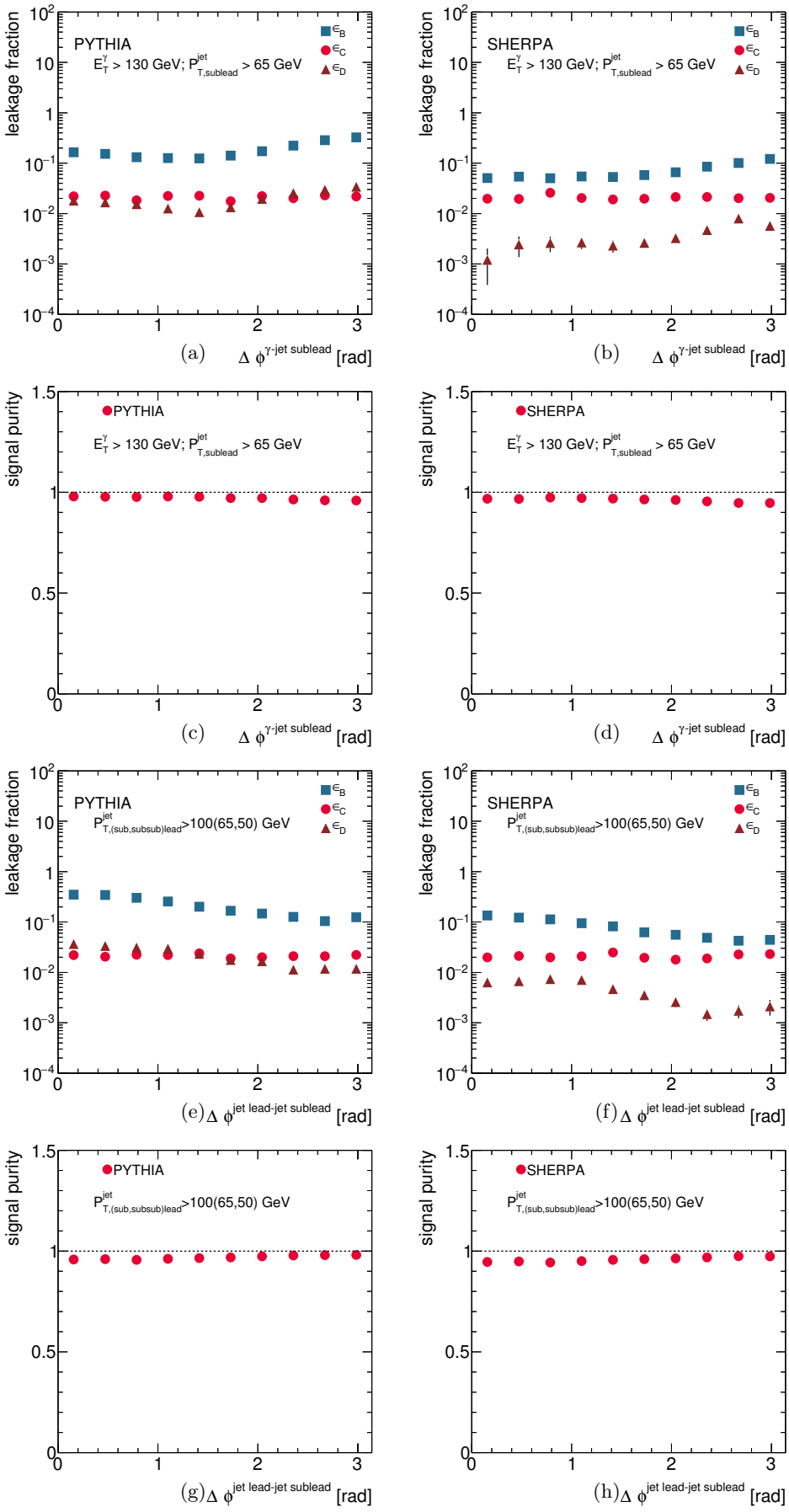


Figure 6.25: *Signal leakage fractions from* PYTHIA *(a,e) and* SHERPA *(b,f) for the B (squares), C (dots) and D (triangles) control regions as functions of* $\Delta\phi^{\gamma-\text{jet sublead}}$ *(a,b) and* $\Delta\phi^{\text{jet lead}-\text{jet sublead}}$ *(e,f) for the photon+two-jets sample. Estimated signal purities in data using signal leakage fractions from* PYTHIA *(c,g) and* SHERPA *(d,h) as functions of* $\Delta\phi^{\gamma-\text{jet sublead}}$ *(c,d) and* $\Delta\phi^{\text{jet lead}-\text{jet sublead}}$ *(g,h) for the photon+two-jets sample. In all plots, the error bars are statistical only.*

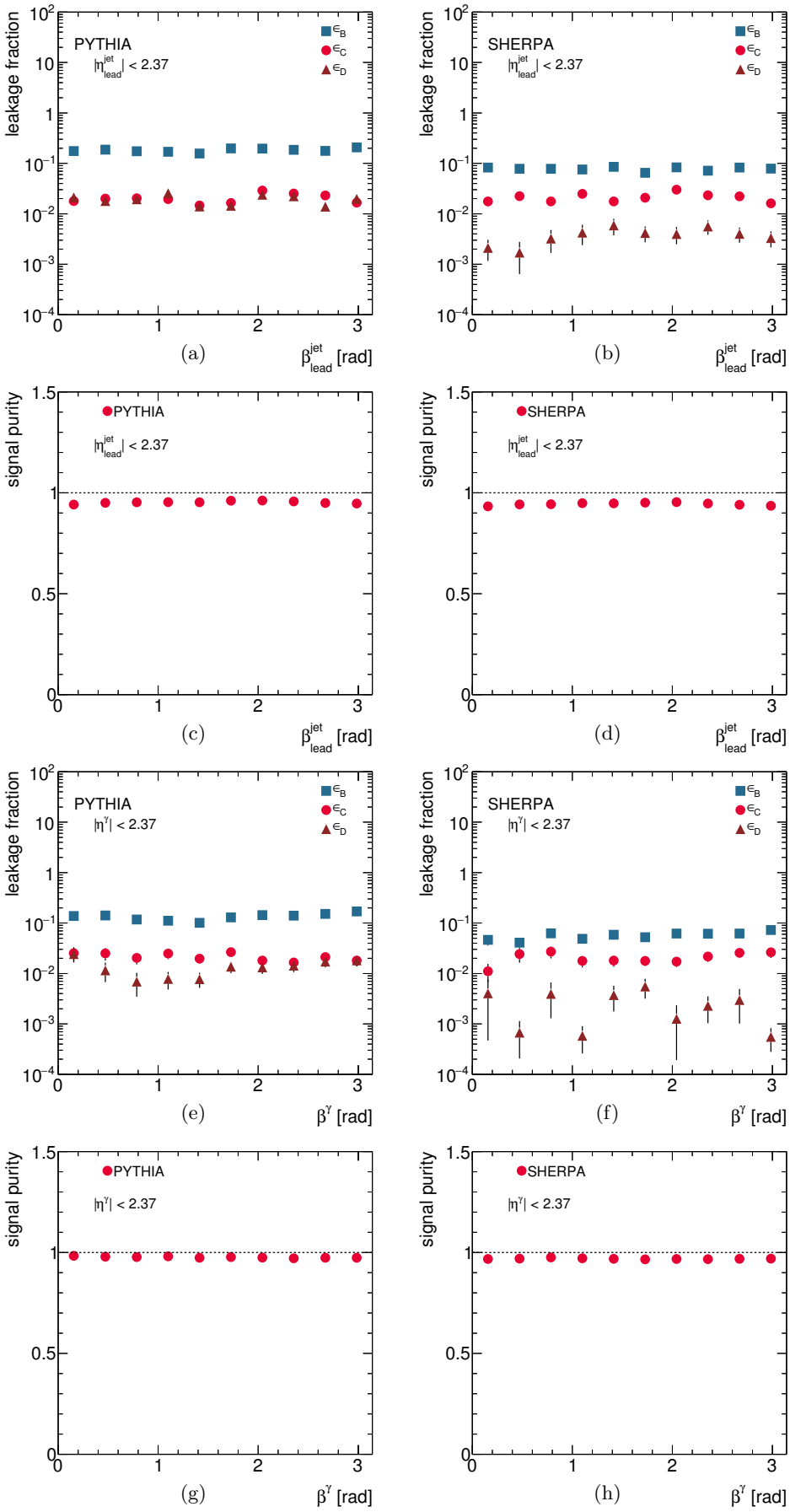


Figure 6.26: *Signal leakage fractions from* PYTHIA *(a,e) and* SHERPA *(b,f) for the B (squares), C (dots) and D (triangles) control regions as functions of* $\beta^{\text{jet}}_{\text{lead}}$ *(a,b) and* β^{γ} *(e,f) for the colour coherence samples. Estimated signal purities in data using signal leakage fractions from* PYTHIA *(c,g) and* SHERPA *(d,h) as functions of* $\beta^{\text{jet}}_{\text{lead}}$ *(c,d) and* β^{γ} *(g,h) for the colour coherence samples. In all plots, the error bars are statistical only.*

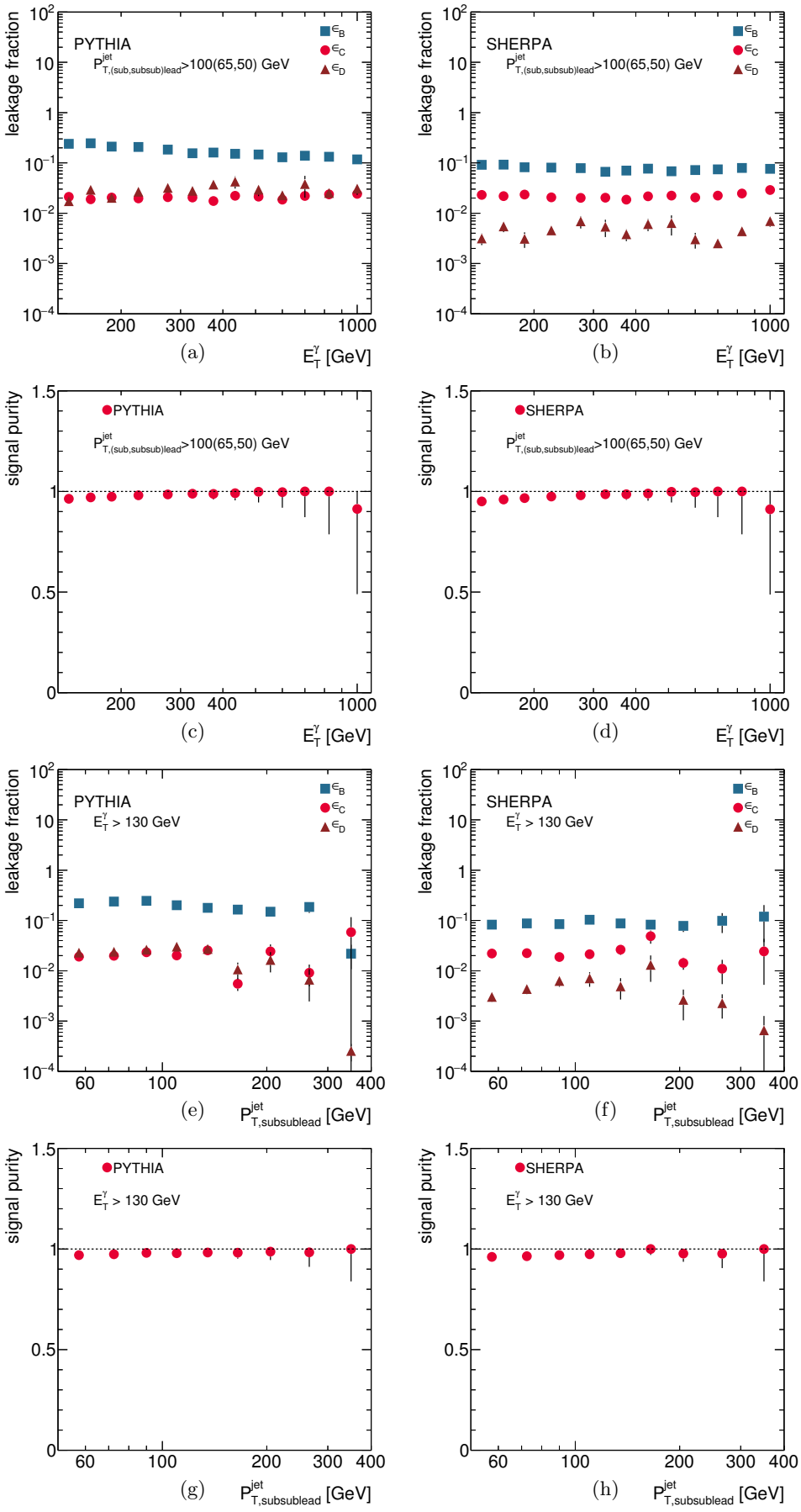


Figure 6.27: *Signal leakage fractions from* PYTHIA *(a,e) and* SHERPA *(b,f) for the B (squares), C (dots) and D (triangles) control regions as functions of* E^{γ}_{T} *(a,b) and* $p^{\mathrm{jet}}_{\mathrm{T,subsublead}}$ *(e,f) for the photon+three-jets sample. Estimated signal purities in data using signal leakage fractions from* PYTHIA *(c,g) and* SHERPA *(d,h) as functions of* E^{γ}_{T} *(c,d) and* $p^{\mathrm{jet}}_{\mathrm{T,subsublead}}$ *(g,h) for the photon+three-jets sample. In all plots, the error bars are statistical only.*

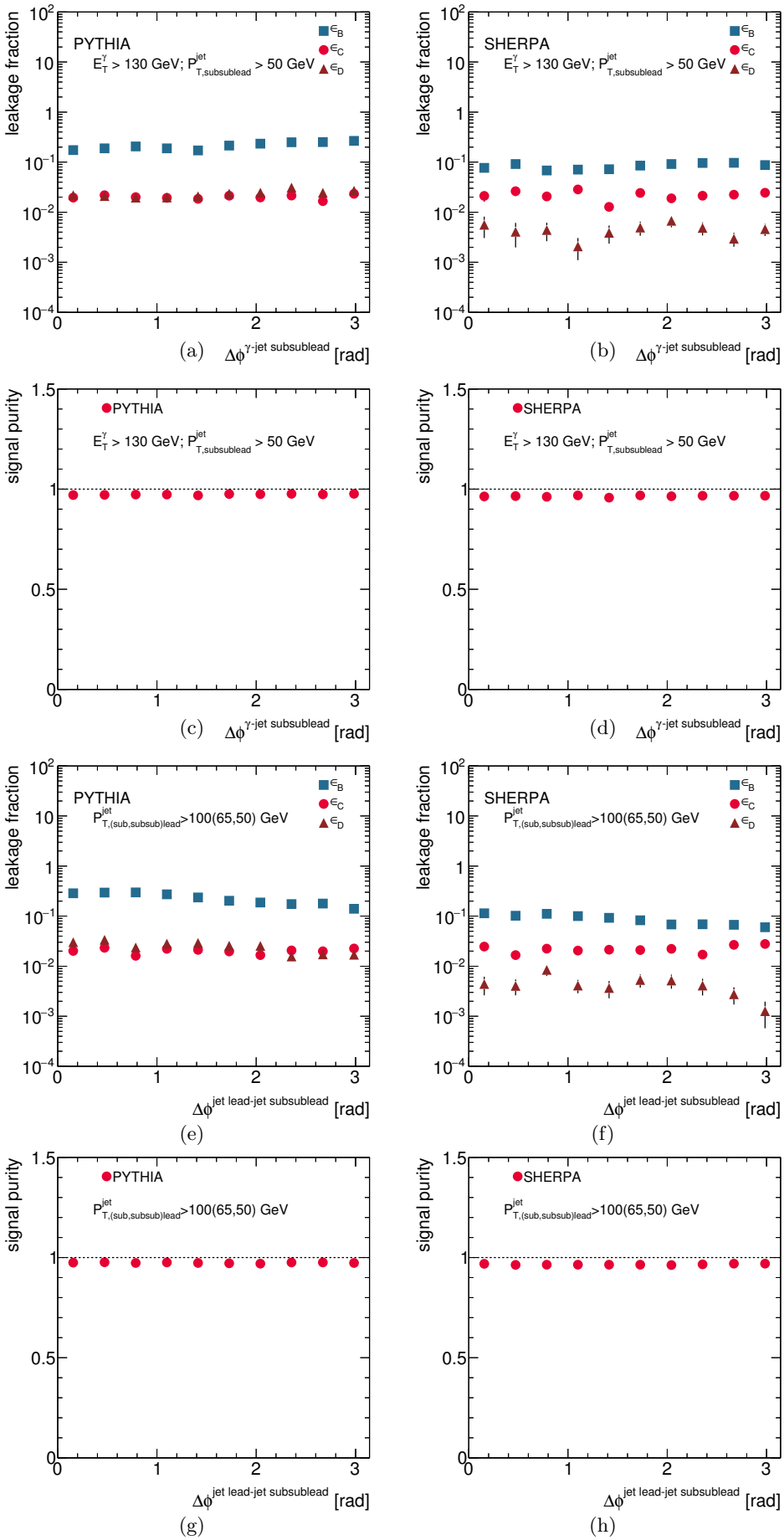


Figure 6.28: *Signal leakage fractions from* PYTHIA *(a,e) and* SHERPA *(b,f) for the B (squares), C (dots) and D (triangles) control regions as functions of* $\Delta\phi^{\gamma-\text{jet subsublead}}$ *(a,b) and* $\Delta\phi^{\text{jet lead}-\text{jet subsublead}}$ *(e,f) for the photon+three-jets sample. Estimated signal purities in data using signal leakage fractions from* PYTHIA *(c,g) and* SHERPA *(d,h) as functions of* $\Delta\phi^{\gamma-\text{jet subsublead}}$ *(c,d) and* $\Delta\phi^{\text{jet lead}-\text{jet sublead}}$ *(g,h) for the photon+three-jets sample. In all plots, the error bars are statistical only.*

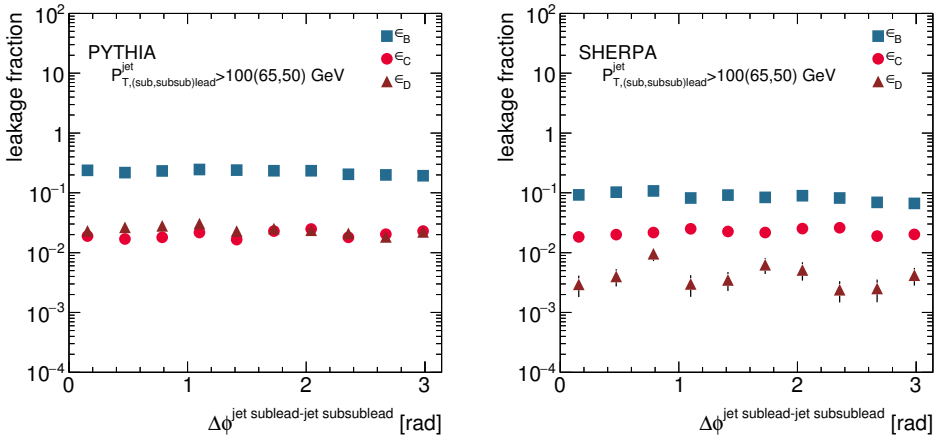


Figure 6.29: *Signal leakage fractions from* PYTHIA *(a) and* SHERPA *(b) for the B (squares), C (dots) and D (triangles) control regions as functions of* $\Delta\phi^{\text{jet sublead}-\text{jet subsublead}}$ *for the photon+three-jets sample. Estimated signal purities in data using signal leakage fractions from* PYTHIA *(c) and* SHERPA *(d) as functions of* $\Delta\phi^{\text{jet sublead}-\text{jet subsublead}}$ *for the photon+three-jets sample. In all plots, the error bars are statistical only.*

6.7 Control plots

In this section, the estimated signal yields using the signal leakage fractions from Pythia or Sherpa are shown.

Photon+one-jet observables

Figure 6.30a and 6.30b show the signal yield as a function of E_{T}^{γ} for the photon+one-jet sample; both Pythia and Sherpa provide a good description of the data, except at high E_{T}^{γ}.

The signal yield as a function of $p_{\mathrm{T,lead}}^{\mathrm{jet}}$ is shown in figures 6.30c and 6.30d; Pythia describes the data only at low $p_{\mathrm{T,lead}}^{\mathrm{jet}}$, whereas Sherpa gives a good description in the full range.

The signal yields as a function of $m^{\gamma j}$ and $|\cos\theta^{\gamma j}|$ are shown in figure 6.31. Both Pythia and Sherpa provide a good description, except at high $m^{\gamma j}$.

Photon+two-jets observables

Figures 6.32a and 6.32b show the signal yield as a function of E_{T}^{γ} for the photon+two-jets sample; both Pythia and Sherpa provide a good description except at high E_{T}^{γ}.

The signal yield as a function of $p_{\mathrm{T,sublead}}^{\mathrm{jet}}$ is shown in figures 6.32c and 6.32d; Pythia describes the data only at low $p_{\mathrm{T,sublead}}^{\mathrm{jet}}$ while Sherpa provides a good description for the full range.

The signal yield as a function of $\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$ and $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$ are shown in figure 6.33; Sherpa provides a good description while Pythia fails to describe the data.

Colour coherence observables

Figure 6.34 show the signal yield as a function of β^{γ} and $\beta_{\mathrm{lead}}^{\mathrm{jet}}$. The simulation of Sherpa gives a good description of the data whereas the description of Pythia is poorer.

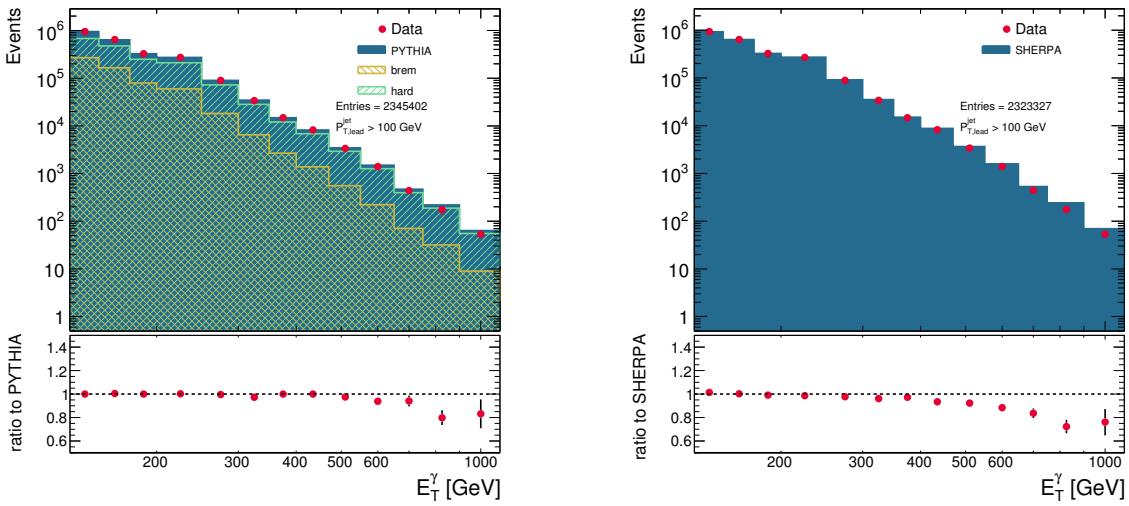
Photon+three-jets observables

Figures 6.35a and 6.35b show the signal yield as a function of E_{T}^{γ} for the photon+three-jets sample; both Pythia and Sherpa provide a good description of the data, except at high E_{T}^{γ}.

The signal yield as a function of $p_{\mathrm{T,subsublead}}^{\mathrm{jet}}$ is shown in figures 6.35c and 6.35d; Pythia describes the data only at low $p_{\mathrm{T,subsublead}}^{\mathrm{jet}}$, whereas Sherpa gives a good description in the full range.

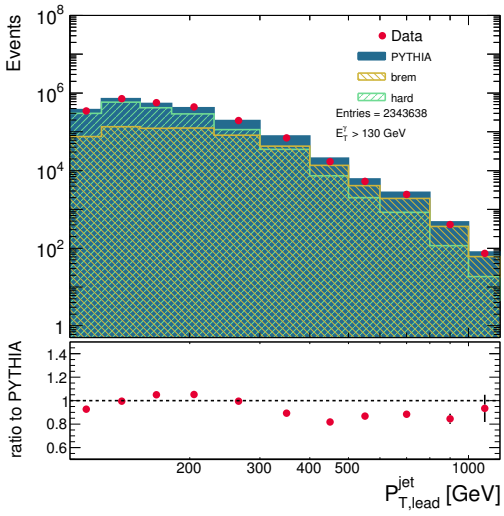
The signal yields as a function of $\Delta\phi^{\gamma-\mathrm{jet\,subsublead}}$ and $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,subsublead}}$ are shown in figure 6.36; Sherpa provides a good description of the data while the description by Pythia is poorer.

The signal yield as a function of $\Delta\phi^{\mathrm{jet\,sublead}-\mathrm{jet\,subsublead}}$ is also shown in figure 6.36; both Pythia and Sherpa provide a good description of the data.

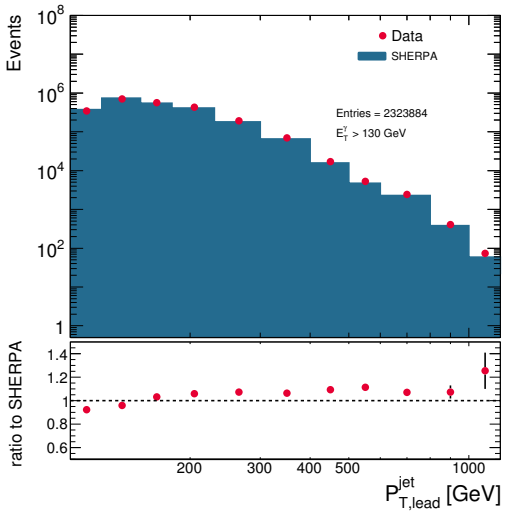


(a)

(b)



(c)



(d)

Figure 6.30: *The estimated signal yields in data (dots) using signal leakage fractions from* PYTHIA *(a,c) or* SHERPA *(b,d) as functions of (a,b) E^{γ}_{T} and (c,d) $p^{\mathrm{jet}}_{\mathrm{T,lead}}$ for the photon+one-jet sample. For comparison, the MC simulations of the signal from* PYTHIA *(a) and* SHERPA *(b) (solid histograms) are also included. The MC distributions are normalised to the number of data events. The hard (right-hatched histograms) and brem (left-hatched histograms) components of* PYTHIA *are also shown in (a,c) and mixed according to the MC prediction. The ratio of the MC to the data is shown in the lower part of the figures.*

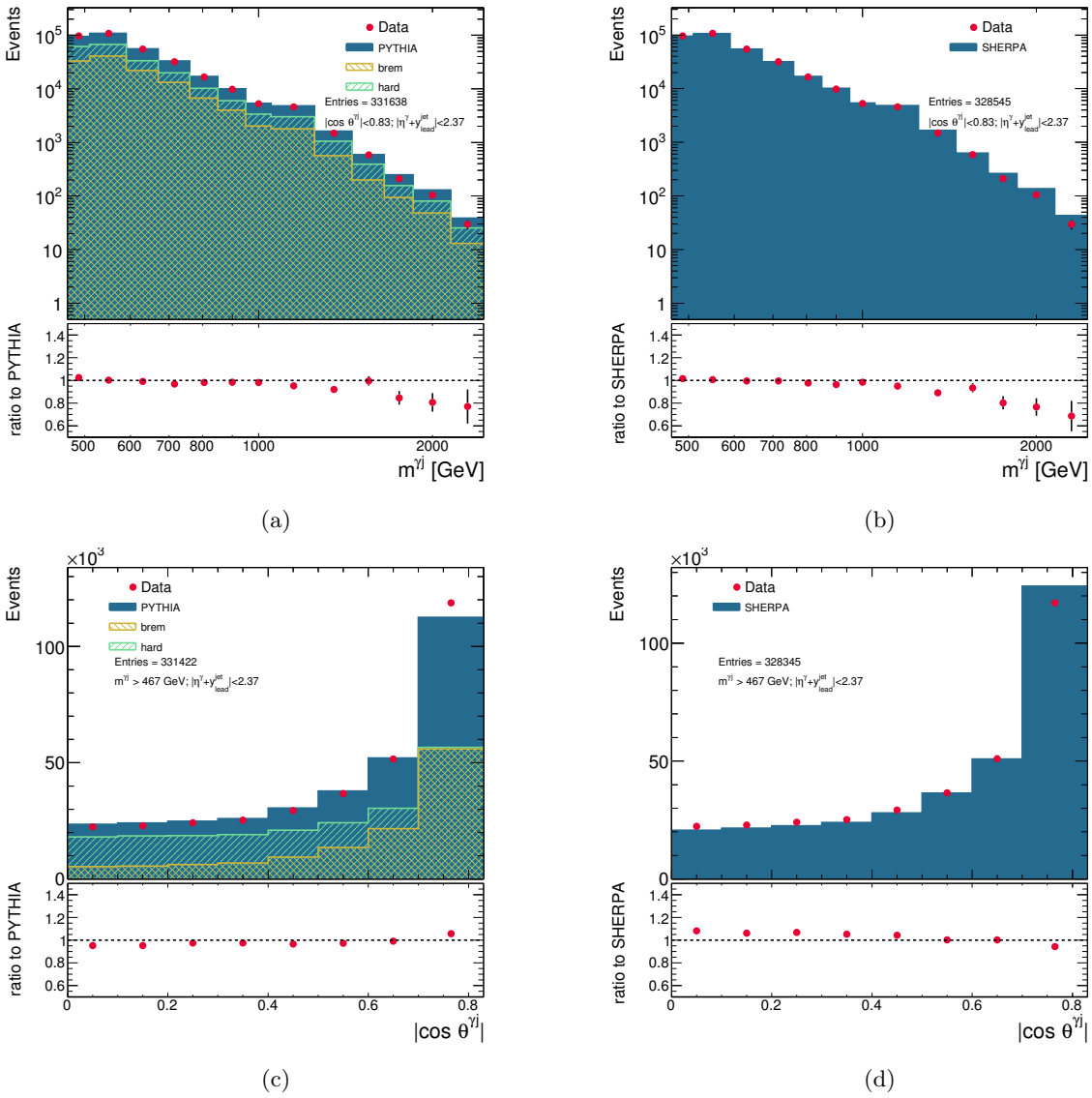


Figure 6.31: *The estimated signal yields in data (dots) using signal leakage fractions from* PYTHIA *(a,c) or* SHERPA *(b,d) as functions of $m^{\gamma j}$ (a,b) and $|\cos\theta^{\gamma j}|$ (c,d) for the photon+one-jet sample. Other details as in the caption to figure 6.30*

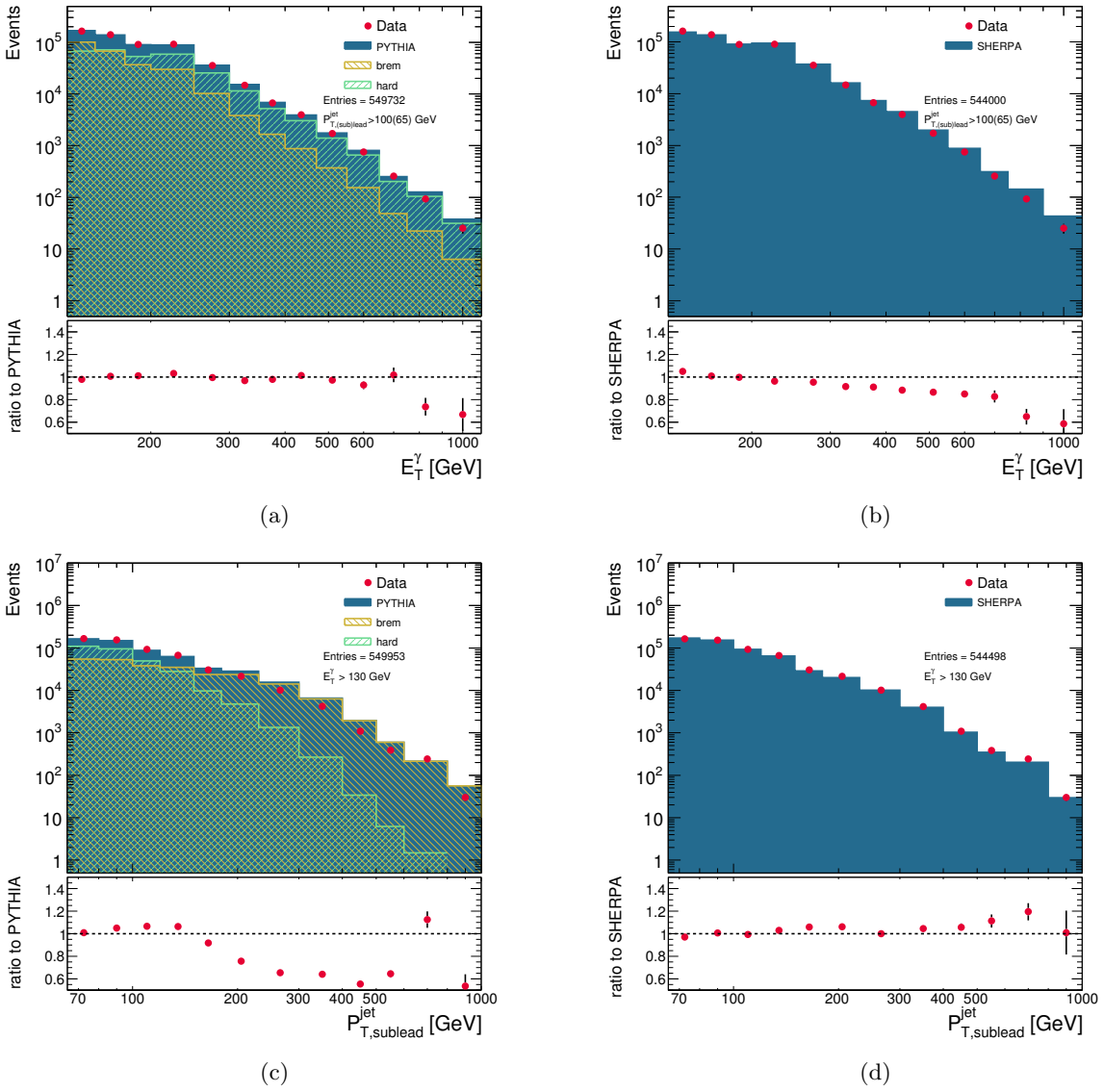


Figure 6.32: *The estimated signal yields in data (dots) using signal leakage fractions from* PYTHIA *(a,c) or* SHERPA *(b,d) as functions of* E_{T}^{γ} *(a,b) and* $p_{\mathrm{T,sublead}}^{\mathrm{jet}}$ *(c,d) for the photon+two-jets sample. Other details as in the caption to figure 6.30*

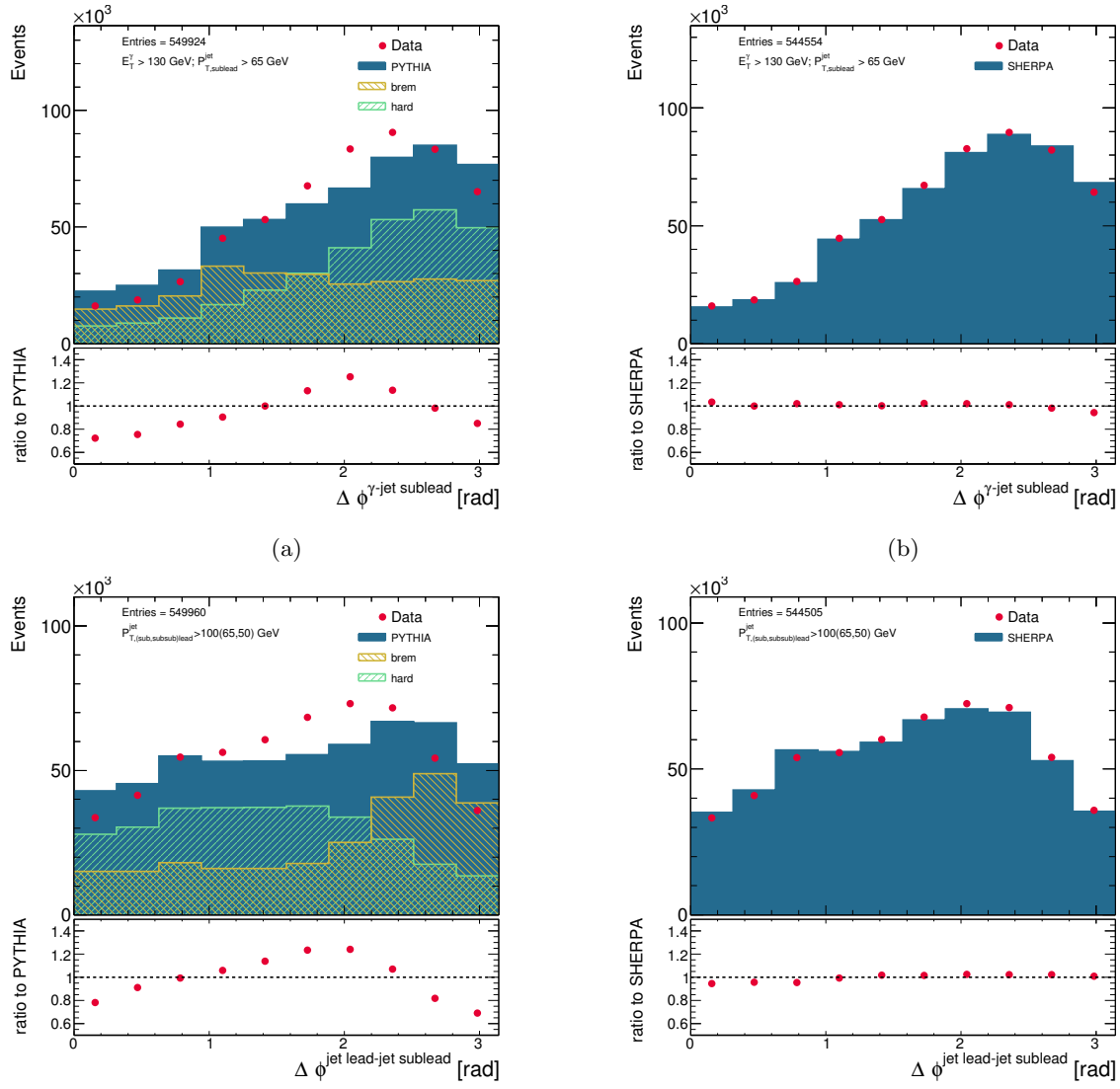
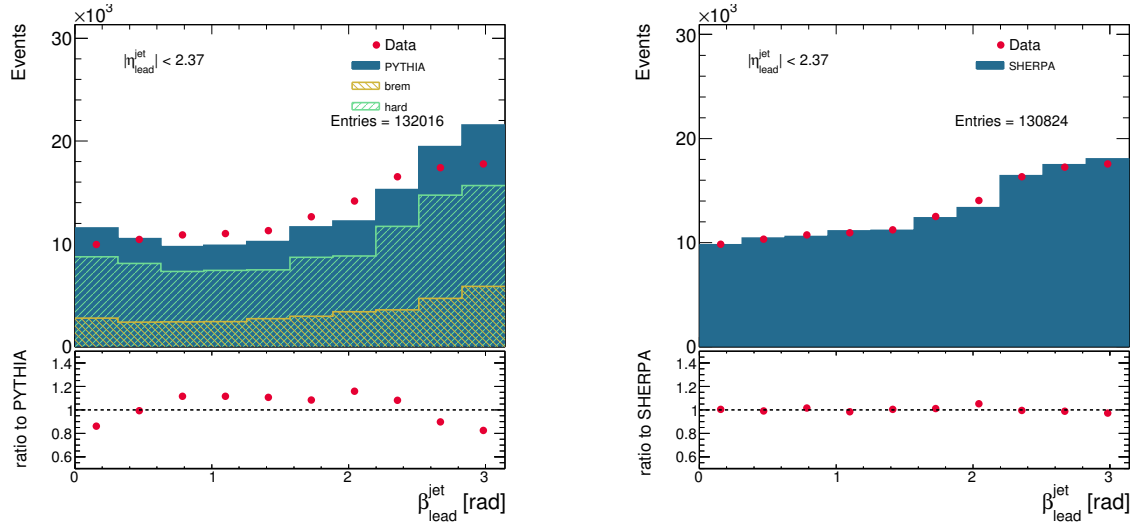
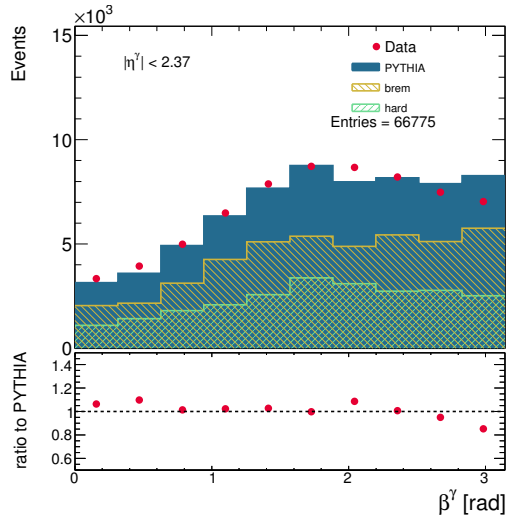


Figure 6.33: *The estimated signal yields in data (dots) using signal leakage fractions from* PYTHIA *(a,c) or* SHERPA *(b,d) as functions of* $\Delta\phi^{\gamma-\text{jet sublead}}$ *(a,b) and* $\Delta\phi^{\text{jet lead}-\text{jet sublead}}$ *(c,d) for the photon+two-jets sample. Other details as in the caption to figure 6.30*

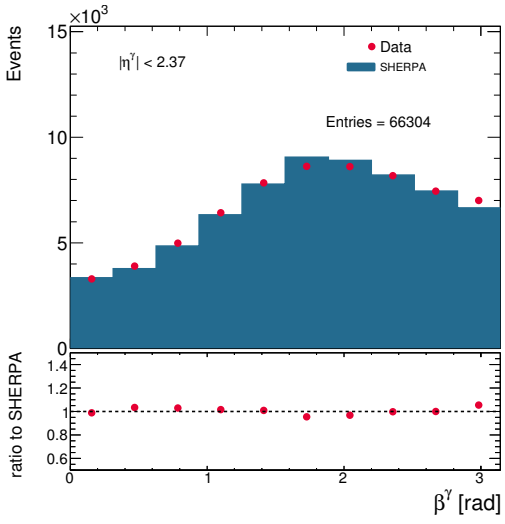


(a)

(b)



(c)



(d)

Figure 6.34: *The estimated signal yields in data (dots) using signal leakage fractions from* PYTHIA *(a,c) or* SHERPA *(b,d) as functions of* $\beta^{\text{jet}}_{\text{lead}}$ *(a,b) and* β^{γ} *(c,d) for the colour coherence samples. Other details as in the caption to figure 6.30*

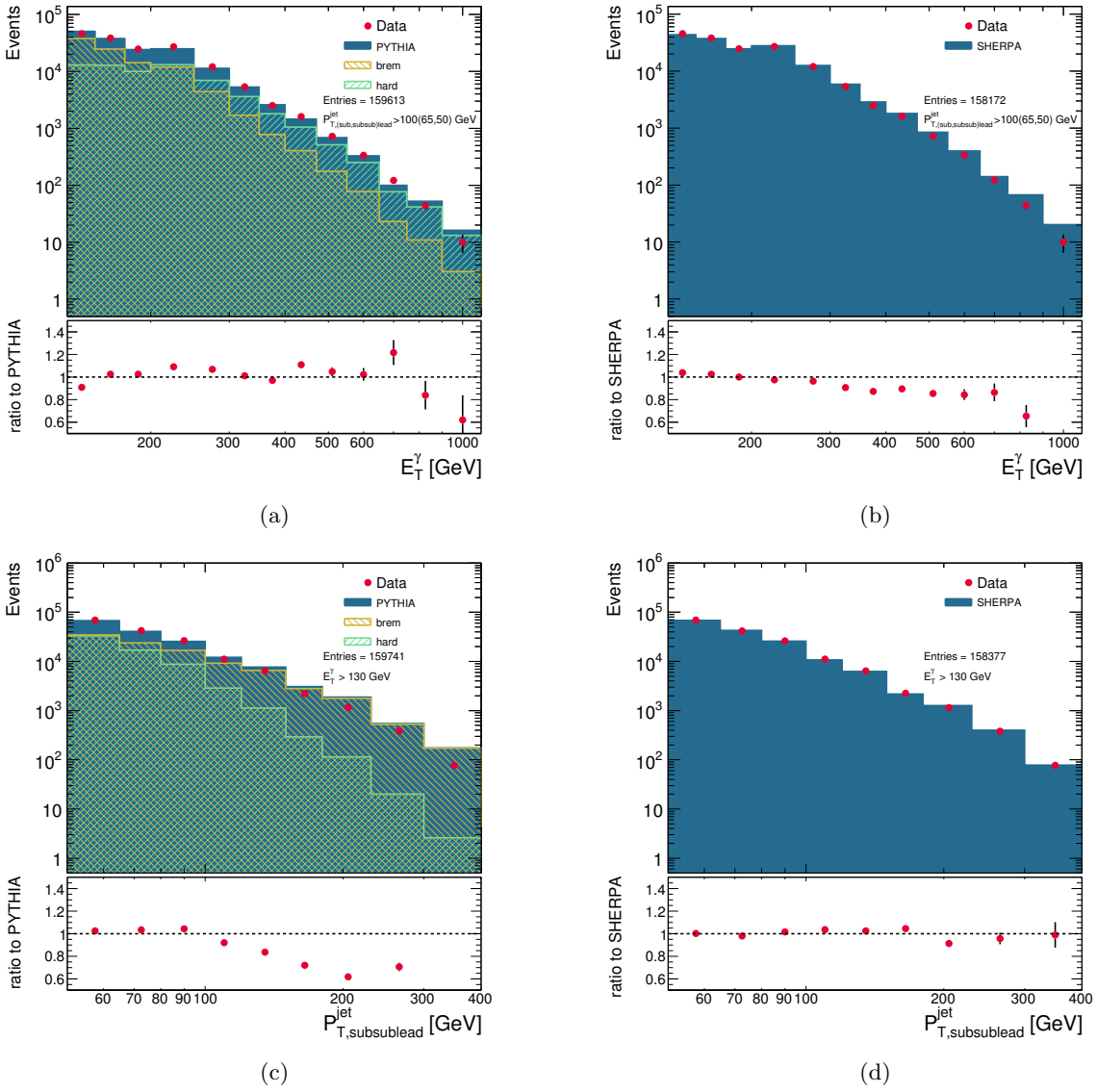


Figure 6.35: *The estimated signal yields in data (dots) using signal leakage fractions from* PYTHIA *(a,c) or* SHERPA *(b,d) as functions of* E_{T}^{γ} *(a,b) and* $p_{\mathrm{T,subsublead}}^{\mathrm{jet}}$ *(c,d) for the photon+three-jets sample. Other details as in the caption to figure 6.30*

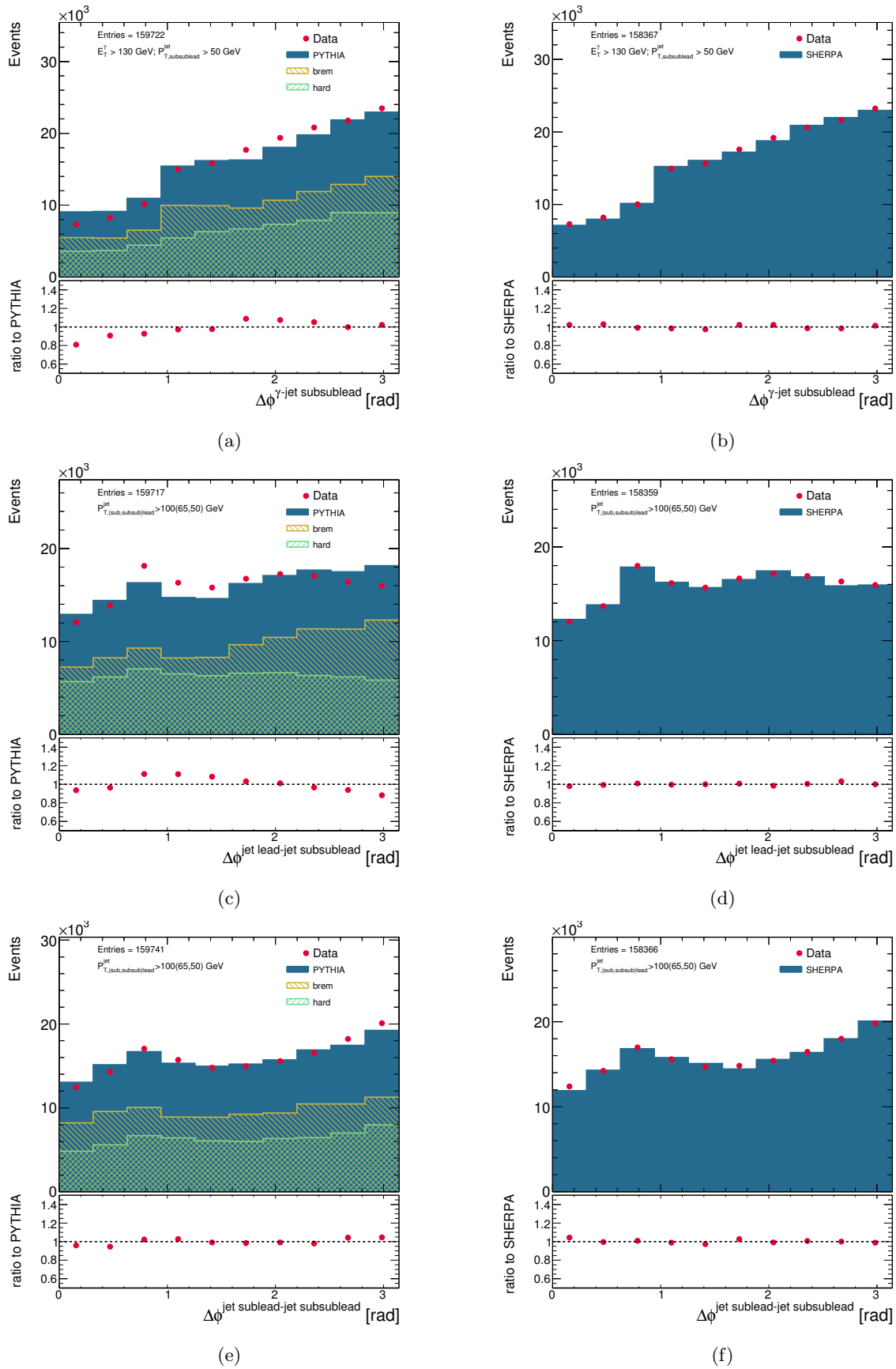


Figure 6.36: *The estimated signal yields in data (dots) using signal leakage fractions from* PYTHIA *(a,c,e) or* SHERPA *(b,d,f) as functions of* $\Delta\phi^{\gamma-\text{jet subsublead}}$ *(a,b),* $\Delta\phi^{\text{jet lead}-\text{jet subsublead}}$ *(c,d) and* $\Delta\phi^{\text{jet sublead}-\text{jet subsublead}}$ *(e,f) for the photon+three-jets sample. Other details as in the caption to figure 6.30*

6.8 Optimisation of the MC description

It is observed in figures 6.30 to 6.36 that the individual components of PYTHIA, namely the hard and brem contributions, have a different shape for several of the studied distributions. Therefore, the shape of the total PYTHIA distributions depend on the relative fraction of the two contributions.

An improvement of the description of the data by the PYTHIA MC was achieved by performing a χ^2 fit to each data distribution of the weight of the hard (α) and the brem ($1-\alpha$) contributions, where α is the free parameter of the fit. The χ^2 function used is

$$\chi^2(\alpha) = \sum_i \left(\frac{N_A^{\mathrm{sig}}(i) - N_A^{\mathrm{MC}}(i,\alpha)}{\Delta N_A^{\mathrm{sig}}(i)} \right)^2 , \tag{6.4}$$

where

$$N_A^{\mathrm{MC}}(i,\alpha) = \frac{N_A^{\mathrm{sig,TOT}}}{\alpha\ N_A^{\mathrm{MC,H,TOT}} + (1-\alpha)\ N_A^{\mathrm{MC,B,TOT}}} \left(\alpha\ N_A^{\mathrm{MC,H}}(i) + \ (1-\alpha)\ N_A^{\mathrm{MC,B}}(i) \right) \tag{6.5}$$

To be consistent, the optimisation of the admixture of the two components should be done simultaneously with the background subtraction since the signal leakage depend on the admixture (see section 6.6). However, such a procedure would result in different estimated signal yield for each fitted variable. To obtain a signal yield independent of the observable, the background subtraction was performed using the default admixture provided by PYTHIA and a systematic uncertainty on the background subtraction due to this admixture was included (see section 6.11).

Figures 6.37 to 6.43 show the same estimated signal yields as figures 6.30 to 6.36 compared to the MC simulations of PYTHIA, after optimising the admixture of the two components, and SHERPA (for SHERPA the figures are exactly the same as before).

An improved description of the data by PYTHIA is obtained, though the description is still poor for some observables, particularly the differences in azimuthal angle for the photon+two-jets sample which are shown in figure 6.40. Figures 6.44 to 6.46 show the χ^2 distributions for each observable.

The fitted values of α are shown in table 6.1, together with the errors in the fit, and are different for each observable. This variation is expected since PYTHIA is only a LO MC; the NLO QCD corrections are expected to affect the two components differently and, furthermore, to entangle them, making any distinction physically impossible. In fact, a variation has been observed [90] in the application of the same procedure at parton-level; the optimal value of α resulting from a fit of the parton-level predictions of the two components in PYTHIA to NLO QCD calculations depends also on the observable. Thus it is understood that the variation of the optimal value of α with the observables arises from higher-order effects; they can be mimicked by mixing the LO descriptions of the two components in an observable-dependent way.

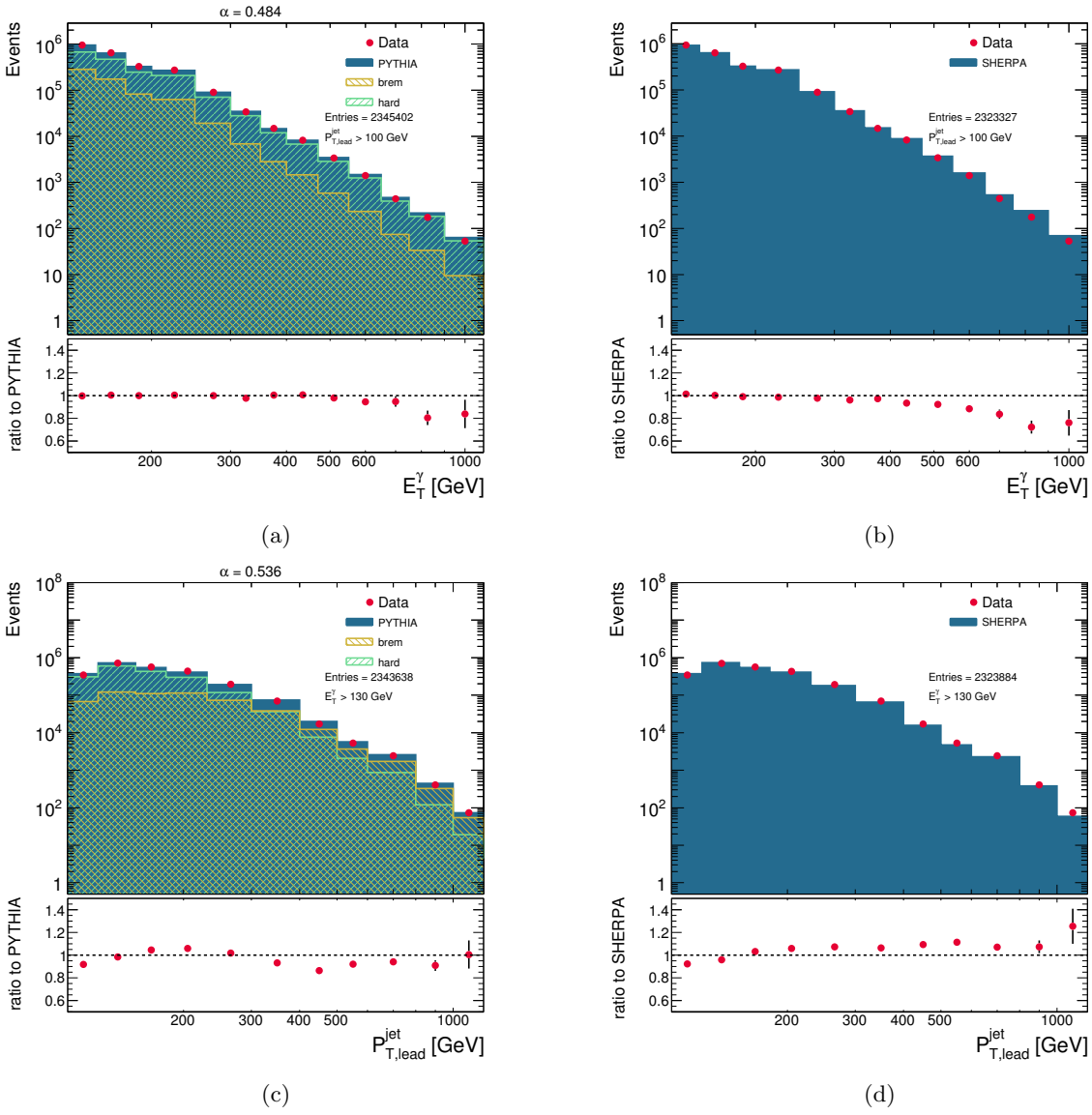


Figure 6.37: *The estimated signal yields in data (dots) using signal leakage fractions from* PYTHIA *(a,c) or* SHERPA *(b,d) as functions of (a,b)* E_{T}^{γ} *and (c,d)* $p_{\mathrm{T,lead}}^{\mathrm{jet}}$ *for the photon+one-jet sample. For comparison, the MC simulations of the signal from* PYTHIA *(a) and* SHERPA *(b) (solid histograms) are also included. The MC distributions are normalised to the number of data events. The hard (right-hatched histograms) and brem (left-hatched histograms) components of* PYTHIA *are also shown in (a,c) and mixed according to the optimised value of* α *shown. The ratio of the MC to the data is shown in the lower part of the figures.*

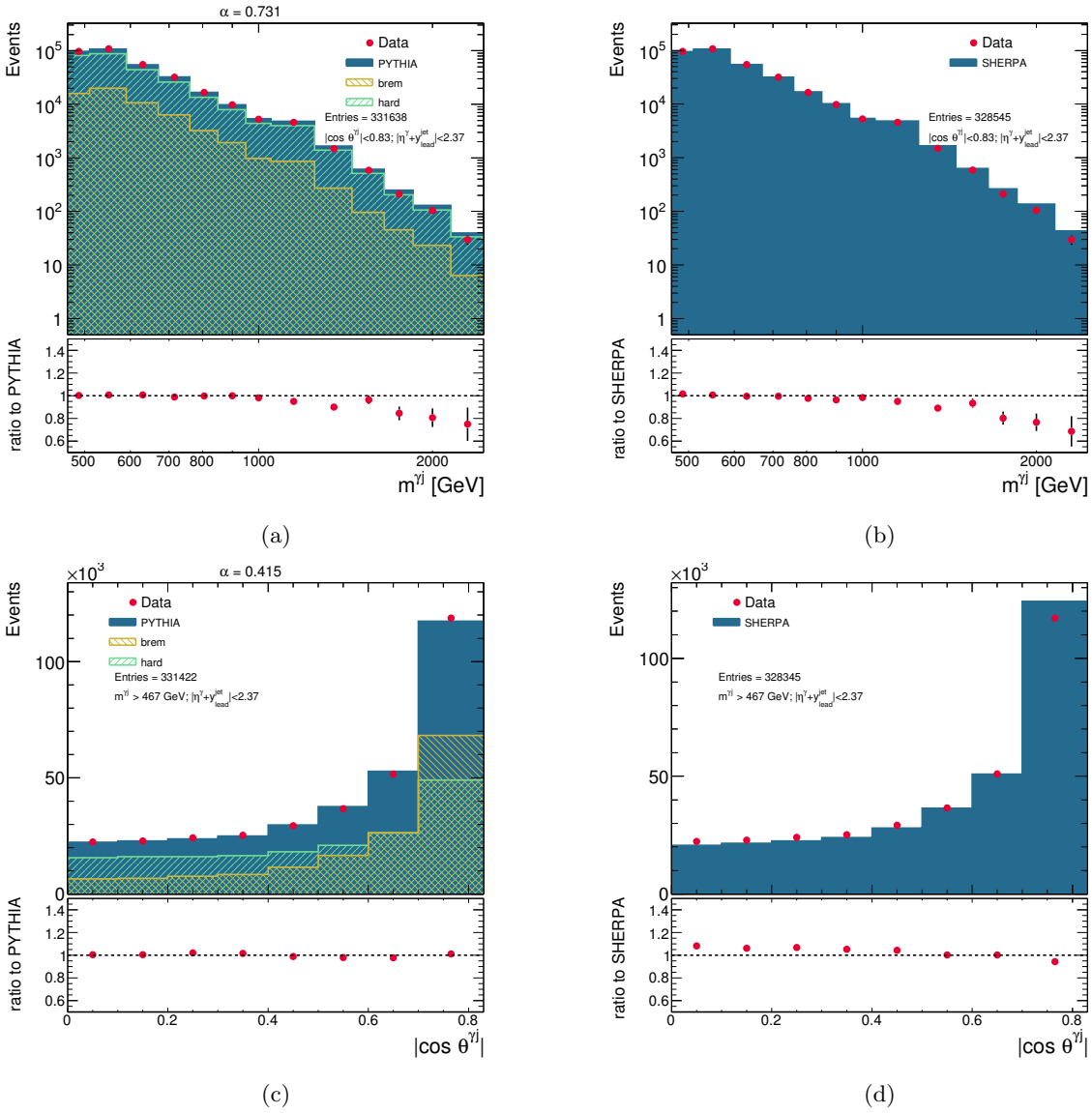


Figure 6.38: *The estimated signal yields in data (dots) using signal leakage fractions from* PYTHIA *(a,c) or* SHERPA *(b,d) as functions of $m^{\gamma j}$ (a,b) and $|\cos\theta^{\gamma j}|$ (c,d) for the photon+one-jet sample. Other details as in the caption to figure 6.37*

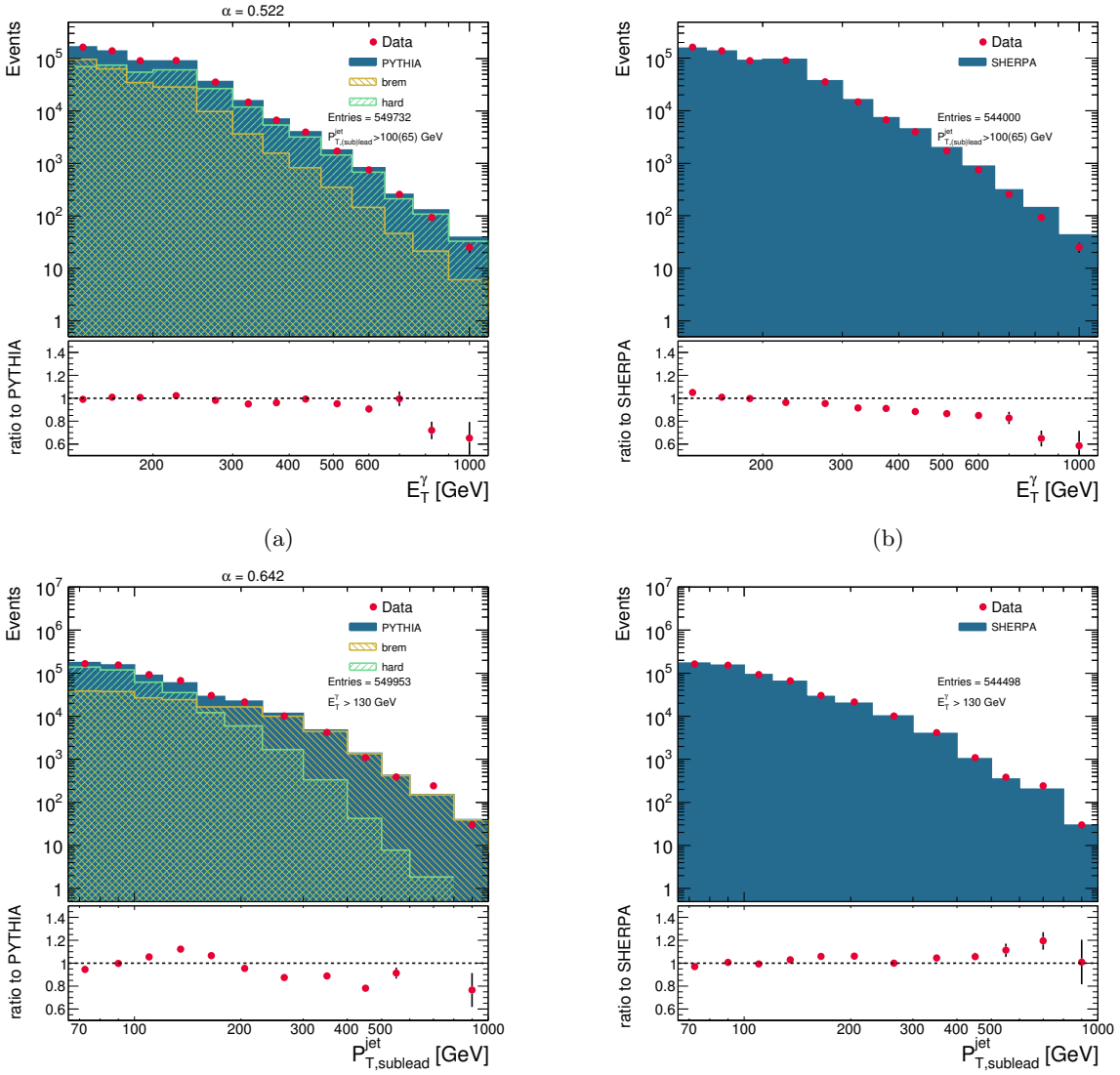


Figure 6.39: *The estimated signal yields in data (dots) using signal leakage fractions from* PYTHIA *(a,c) or* SHERPA *(b,d) as functions of E_{T}^{γ} (a,b) and $p_{\mathrm{T,sublead}}^{\mathrm{jet}}$ (c,d) for the photon+two-jets sample. Other details as in the caption to figure 6.37*

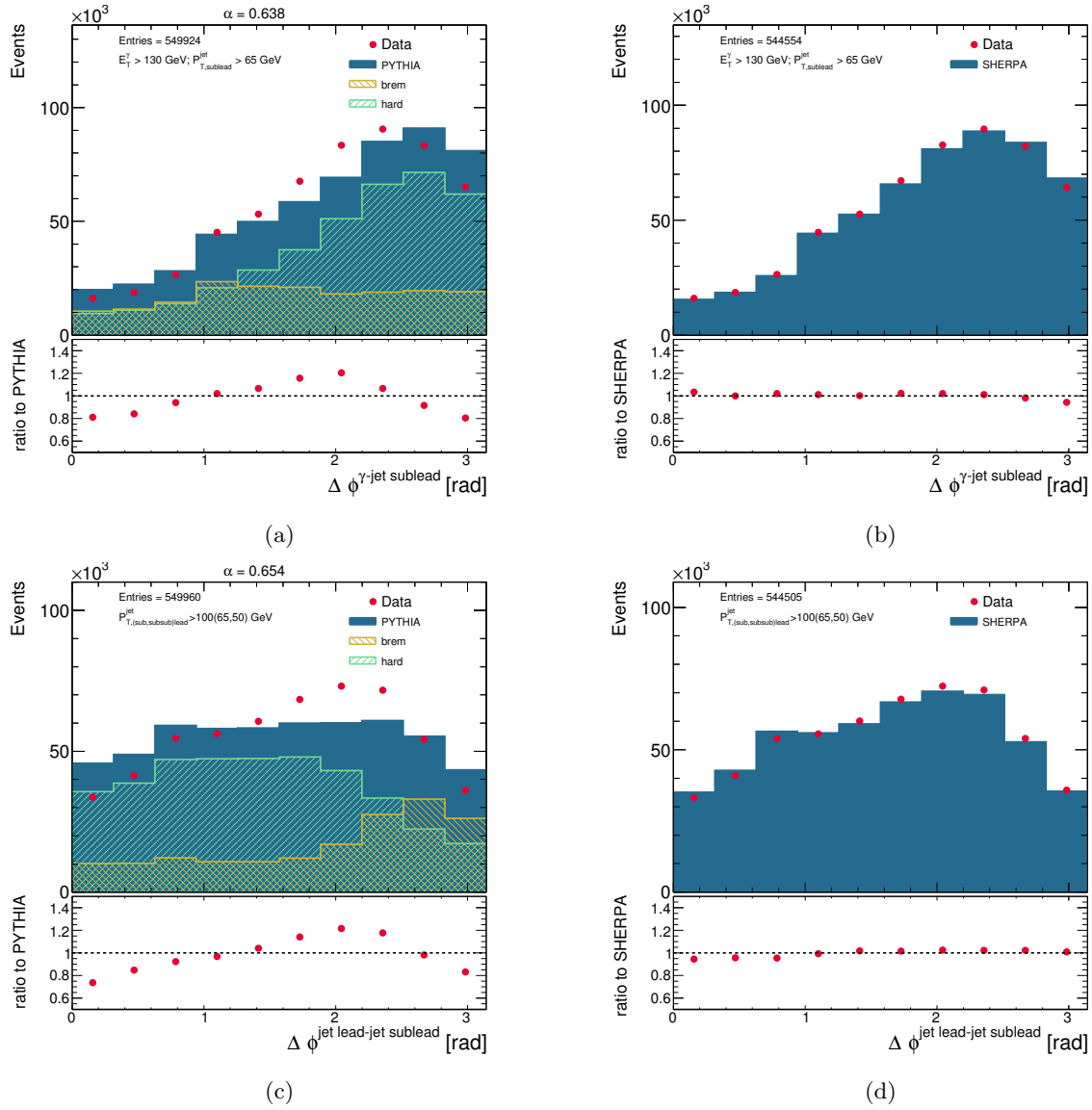
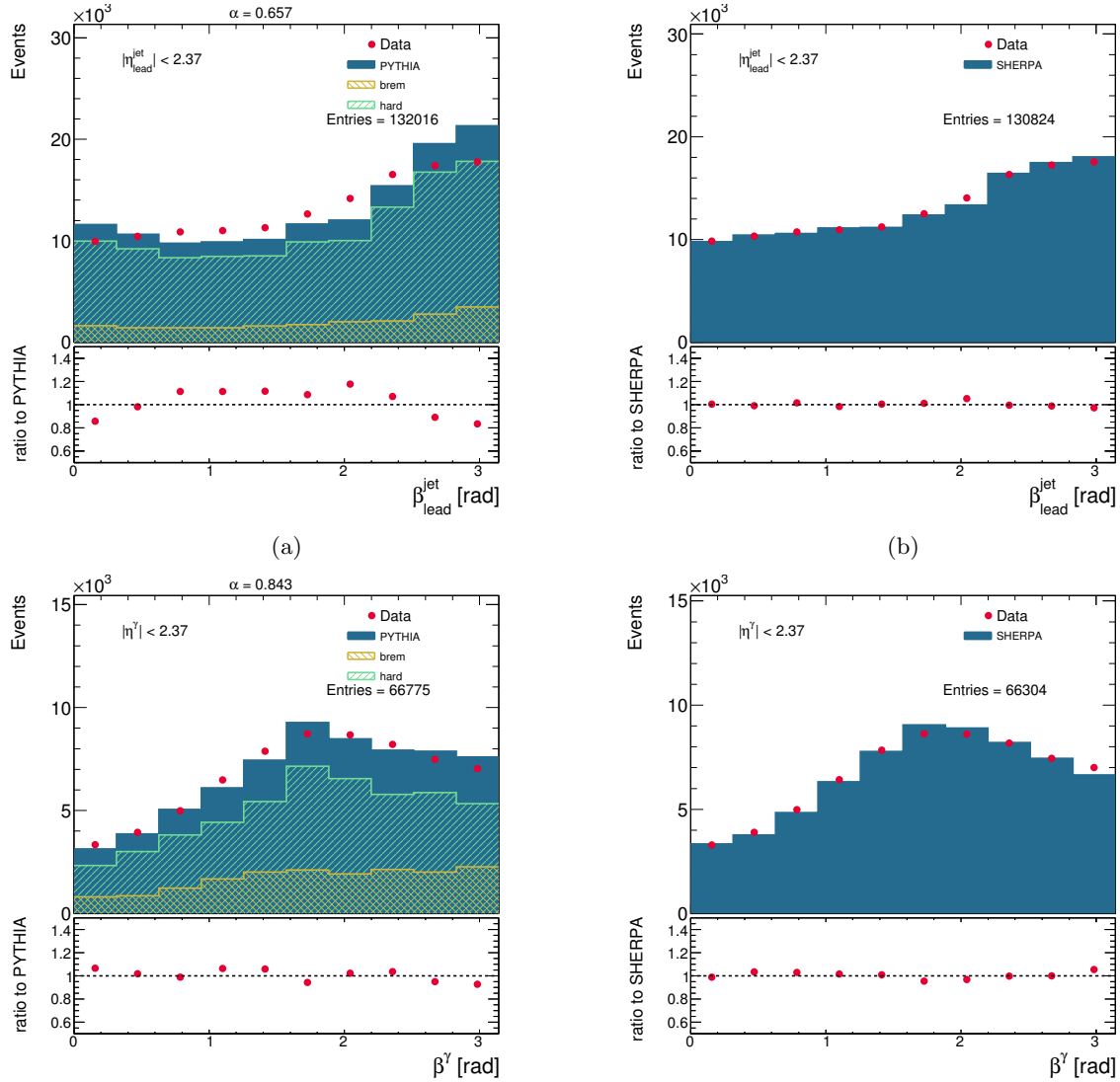


Figure 6.40: *The estimated signal yields in data (dots) using signal leakage fractions from* PYTHIA *(a,c) or* SHERPA *(b,d) as functions of* $\Delta\phi^{\gamma-\text{jet sublead}}$ *(a,b) and* $\Delta\phi^{\text{jet lead}-\text{jet sublead}}$ *(c,d) for the photon+two-jets sample. Other details as in the caption to figure 6.37*



(a) (b) (c) (d)

Figure 6.41: *The estimated signal yields in data (dots) using signal leakage fractions from* PYTHIA *(a,c) or* SHERPA *(b,d) as functions of* $\beta^{\text{jet}}_{\text{lead}}$ *(a,b) and* β^{γ} *(c,d) for the colour coherence samples. Other details as in the caption to figure 6.37*

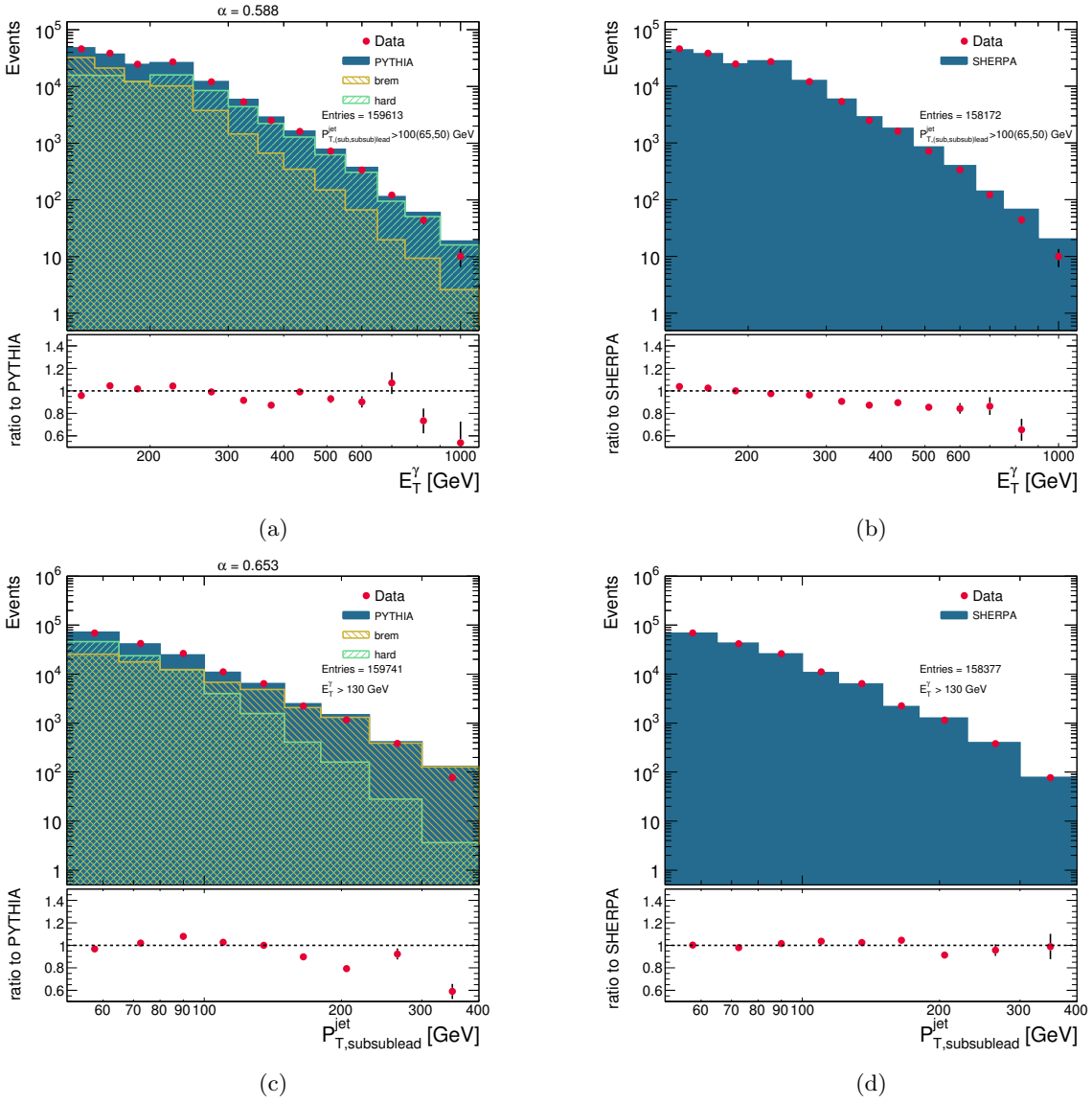


Figure 6.42: *The estimated signal yields in data (dots) using signal leakage fractions from* PYTHIA *(a,c) or* SHERPA *(b,d) as functions of* E_{T}^{γ} *(a,b) and* $p_{\mathrm{T,subsublead}}^{\mathrm{jet}}$ *(c,d) for the photon+three-jets sample. Other details as in the caption to figure 6.37*

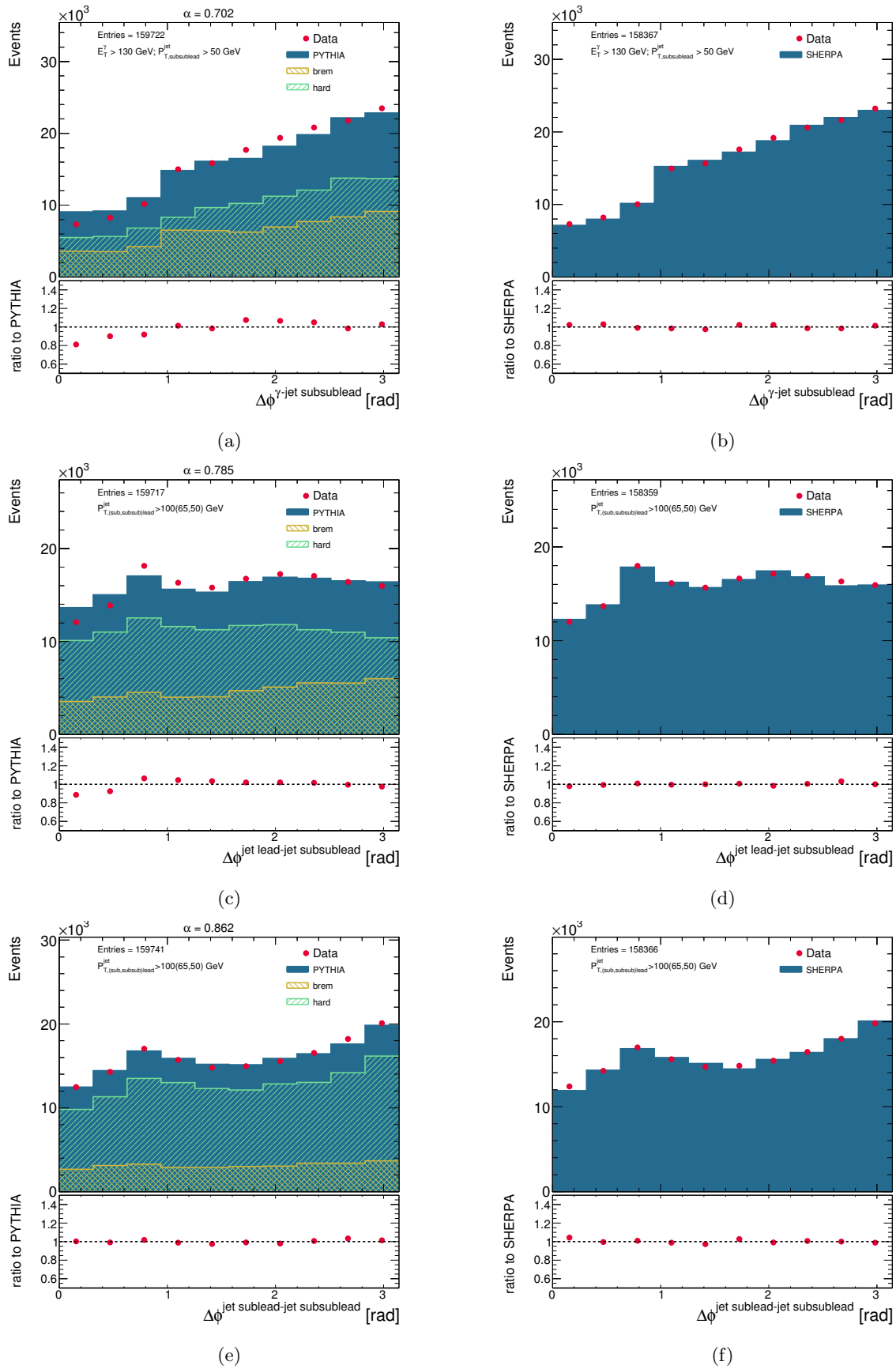


Figure 6.43: *The estimated signal yields in data (dots) using signal leakage fractions from* PYTHIA *(a,c,e) or* SHERPA *(b,d,f) as functions of* $\Delta\phi^{\gamma-\text{jet subsublead}}$ *(a,b),* $\Delta\phi^{\text{jet lead}-\text{jet subsublead}}$ *(c,d) and* $\Delta\phi^{\text{jet sublead}-\text{jet subsublead}}$ *(e,f) for the photon+three-jets sample. Other details as in the caption to figure 6.37*

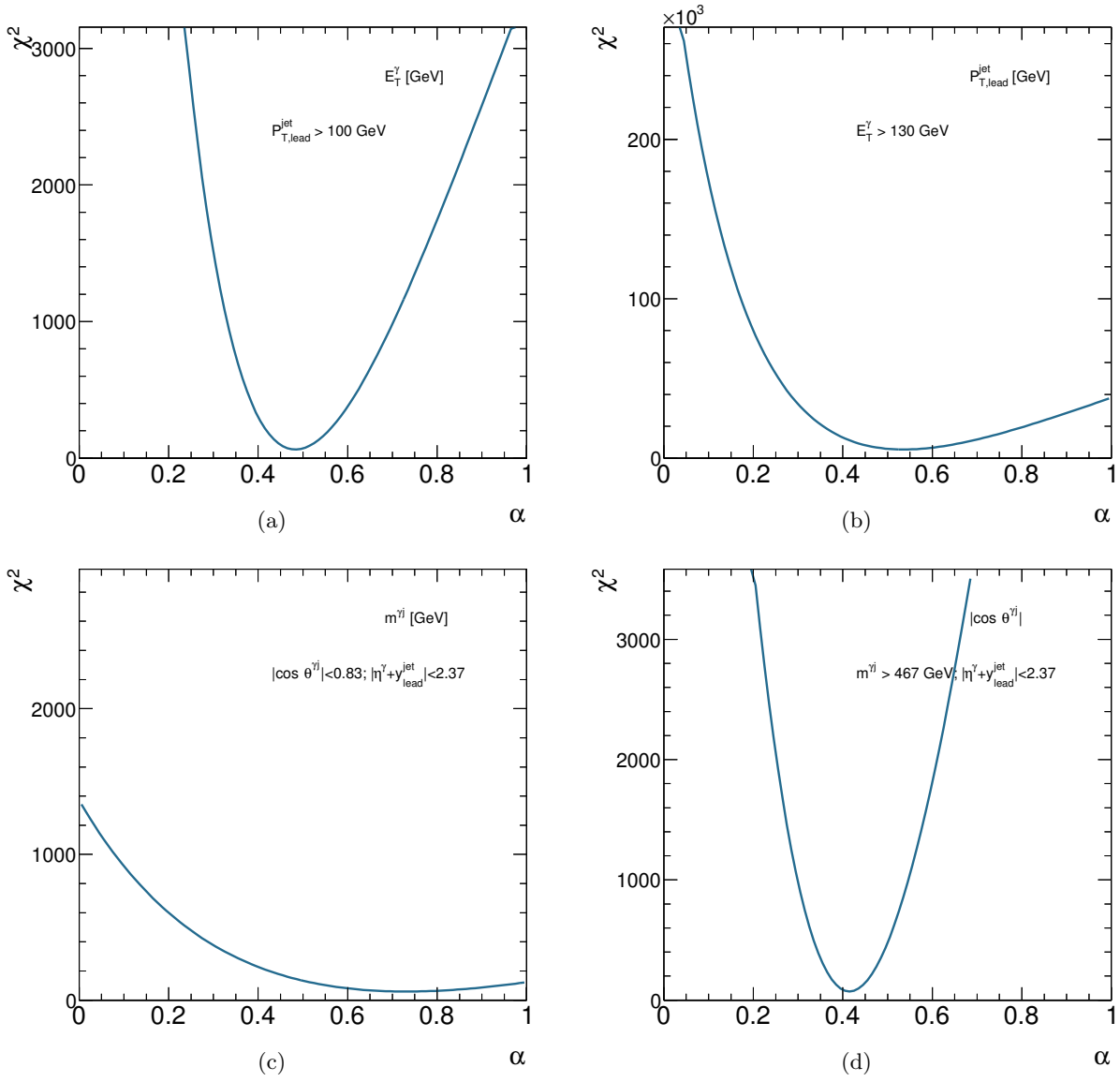


Figure 6.44: *The χ^2 distributions resulting from the optimisation of the admixture of the two* Pythia *components to signal yield distributions as functions of E_{T}^{γ} (a), $p_{\mathrm{T,lead}}^{\mathrm{jet}}$ (b), $m^{\gamma j}$ (c) and $|\cos\theta^{\gamma j}|$ (d) for the photon+one-jet sample.*

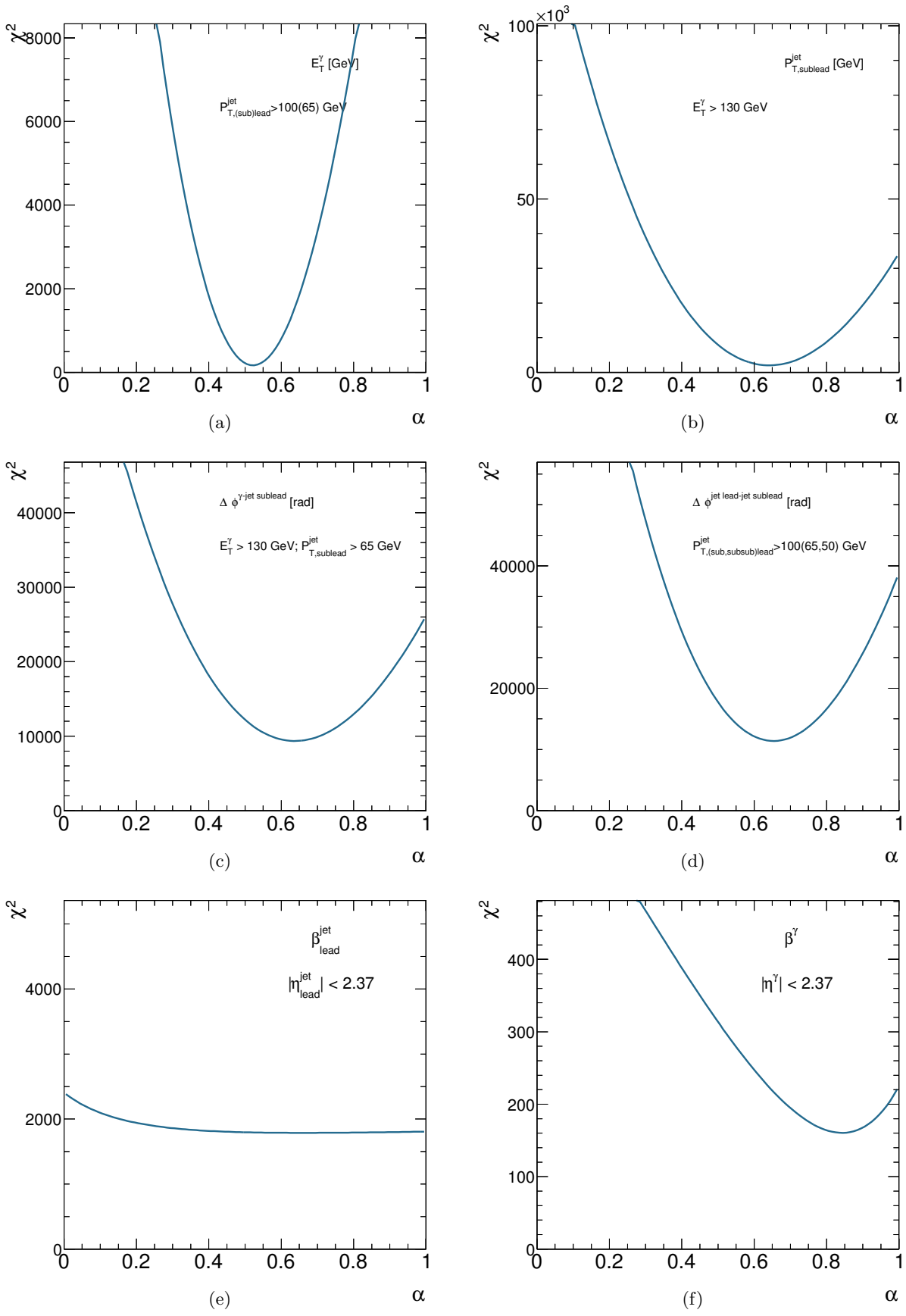


Figure 6.45: *The χ^2 distributions resulting from the optimisation of the admixture of the two* PYTHIA *components to signal yield distributions as functions of E_{T}^{γ} (a), $p_{\mathrm{T,sublead}}^{\mathrm{jet}}$ (b), $\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$ (c), $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$ (d), $\beta_{\mathrm{lead}}^{\mathrm{jet}}$ (e) and β^{γ} (f) for the photon+two-jets (a,b,c,d) and the colour coherence (e,f) samples.*

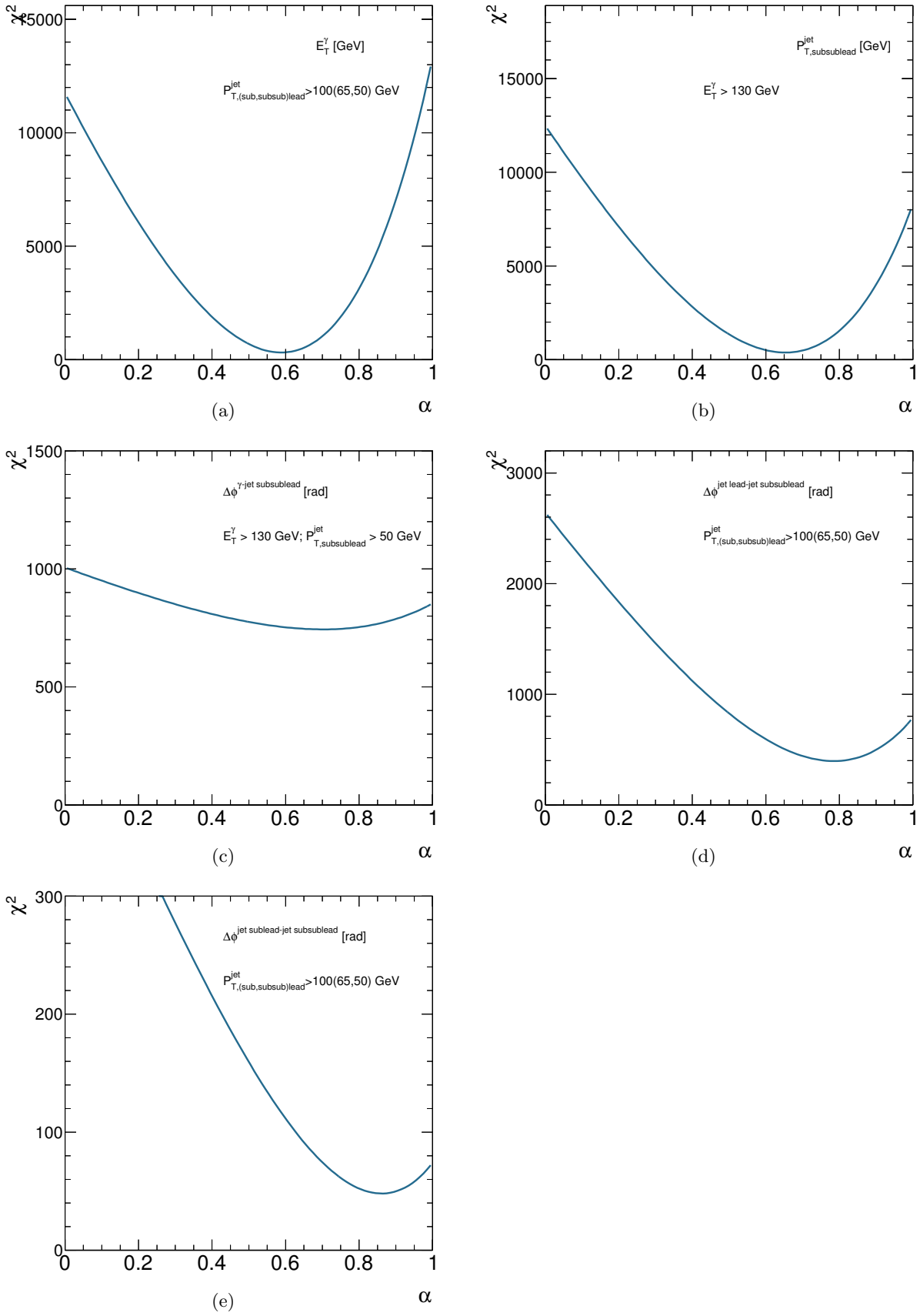


Figure 6.46: *The χ^2 distributions resulting from the optimisation of the admixture of the two* Pythia *components to signal yield distributions as functions of E_{T}^{γ} (a), $p_{\mathrm{T,subsublead}}^{\mathrm{jet}}$ (b), $\Delta\phi^{\gamma-\mathrm{jet\,subsublead}}$ (c), $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,subsublead}}$ (d), $\Delta\phi^{\mathrm{jet\,sublead}-\mathrm{jet\,subsublead}}$ (e) for the photon+three-jets sample.*

	Pythia	
	α	$\pm$
E_{T}^{γ} (one jet)	0.4835	0.0059
$p_{\mathrm{T,lead}}^{\mathrm{jet}}$ (one jet)	0.5358	0.0018
$m^{\gamma j}$ (one jet)	0.731	0.030
$\lvert\cos\theta^{\gamma j}\rvert$ (one jet)	0.4148	0.0041
E_{T}^{γ} (two jets)	0.5219	0.0031
$p_{\mathrm{T,sublead}}^{\mathrm{jet}}$ (two jets)	0.6415	0.0019
$\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$ (two jets)	0.6378	0.0026
$\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$ (two jets)	0.6536	0.0020
$\beta_{\mathrm{lead}}^{\mathrm{jet}}$ (two jets)	0.657	0.059
β^{γ} (two jets)	0.843	0.022
E_{T}^{γ} (three jets)	0.5881	0.0044
$p_{\mathrm{T,subsublead}}^{\mathrm{jet}}$ (three jets)	0.6531	0.0046
$\Delta\phi^{\gamma-\mathrm{jet\,subsublead}}$ (three jets)	0.702	0.033
$\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,subsublead}}$ (three jets)	0.785	0.012
$\Delta\phi^{\mathrm{jet\,sublead}-\mathrm{jet\,subsublead}}$ (three jets)	0.862	0.029

Table 6.1: *Values of the free parameter α and its statistical uncertainty in the admixture of the hard and brem components in* Pythia *resulting from the fit to the data distributions.*

6.9 Selection efficiency and purity

The selection efficiency and purity were evaluated using the MC samples. The integrated selection efficiency was computed as

$$\epsilon = \frac{N^{\mathrm{reco,part}}}{N^{\mathrm{part}}}, \tag{6.6}$$

where $N^{\mathrm{reco,part}}$ is the number of events that pass the selection criteria both at the detector and particle levels and N^{part} is the number of events that pass the selection criteria at the particle level. The integrated selection efficiency was found to be 81.5% (81.3%) for the photon+one-jet sample, 79.4% (78.9%) for the photon+one-jet $m^{\gamma j}$-$\cos\theta^{\gamma j}$ sample, 75.3% (74.6%) for the photon+two-jets sample, 69.8% (69.2%) for the $\beta_{\mathrm{lead}}^{\mathrm{jet}}$ sample, 73.4% (74.5%) for the β^{γ} sample and 70.6% (70.2%) for the photon+three-jets sample using Pythia (Sherpa).

The bin-to-bin selection efficiency was computed as:

$$\epsilon_i = \frac{N_i^{\mathrm{reco,part}}}{N_i^{\mathrm{part}}}, \tag{6.7}$$

where $N_i^{\mathrm{reco,part}}$ is the number of events that pass the selection criteria both at the detector and particle levels and are generated and reconstructed in bin i and N_i^{part} is the number of events that pass the selection criteria at the particle level and are generated in bin i. Figures 6.47 to 6.49 show the bin-to-bin selection efficiencies as functions of the observables studied for both Pythia and Sherpa. The bin-to-bin efficiency is typically above 50% and very similar for both MC generators.

The integrated selection purity was computed as

$$\mathcal{P} = \frac{N^{\text{reco,part}}}{N^{\text{reco}}}, \tag{6.8}$$

where $N^{\text{reco,part}}$ is the number of MC events that pass all the selection requirements both at the detector and particle levels and N^{reco} is the number of MC events that pass the selection requirements at the detector level. The integrated selection purity was found to be 93.9% (94.6%) for the photon+one-jet sample, 91.1% (92.0%) for the photon+one-jet $m^{\gamma j}$-$\cos\theta^{\gamma j}$ sample, 83.4% (84.4%) for the photon+two-jets sample, 74.7% (75.3%) for the $\beta_{\text{lead}}^{\text{jet}}$ sample, 76.0% (79.6%) for the β^{γ} sample and 76.6% (78.2%) for the photon+three-jets sample using Pythia (Sherpa).

The bin-to-bin selection purity was computed as

$$\mathcal{P}_i = \frac{N_i^{\text{reco,part}}}{N_i^{\text{reco}}}, \tag{6.9}$$

where $N_i^{\text{reco,part}}$ is the number of events that pass the selection criteria both at the detector and particle levels and are generated and reconstructed in bin i and N_i^{reco} is the number of events that pass the selection criteria at the detector level and are reconstructed in bin i. Figures 6.50 to 6.52 show the bin-to-bin selection efficiencies as functions of the observables studied for both Pythia and Sherpa. The bin-to-bin purity is typically above 50% and very similar for both MC generators.

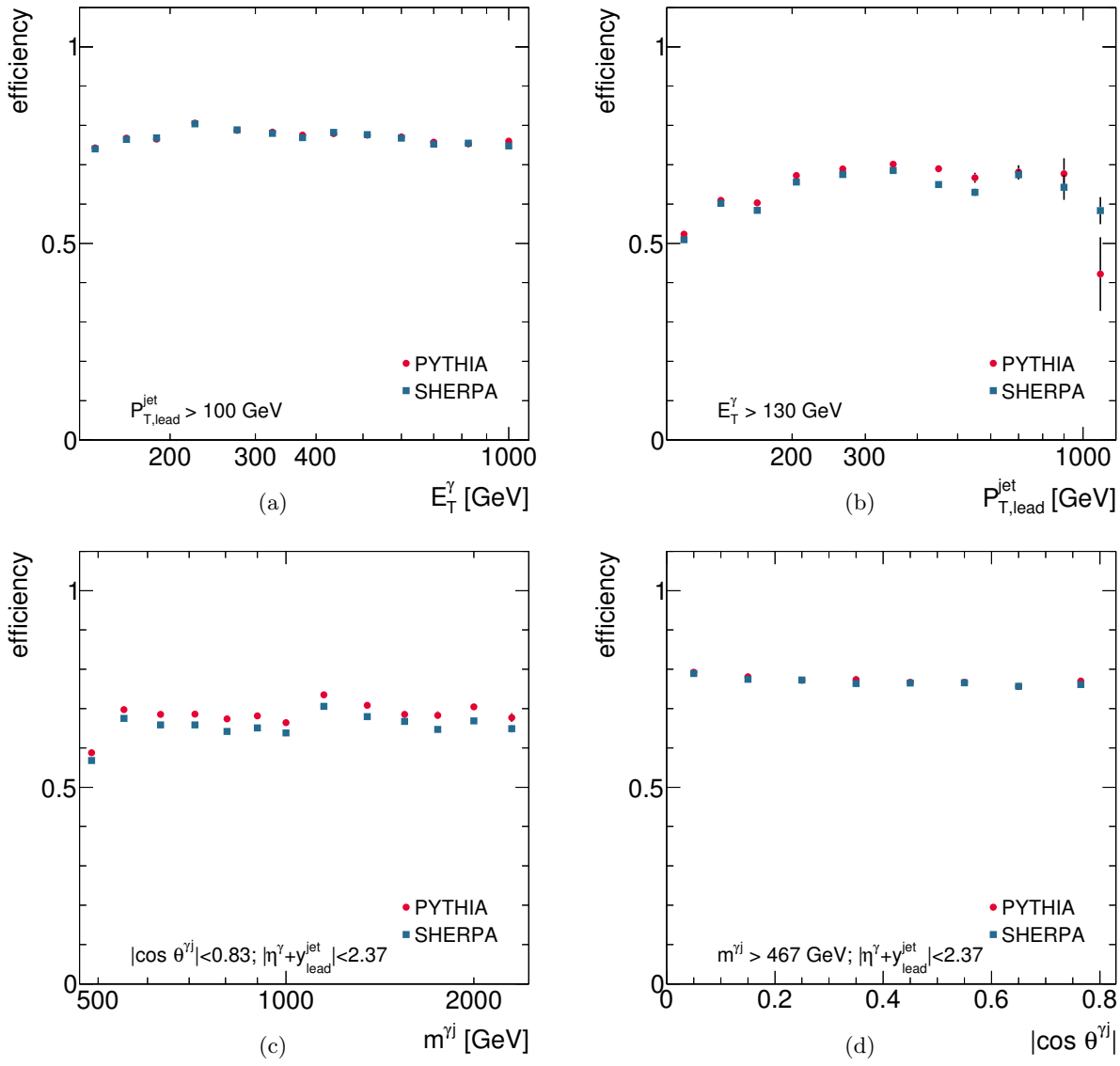


Figure 6.47: *Selection efficiency from* PYTHIA *(dots) and* SHERPA *(squares) as a function of* E^{γ}_{T} *(a),* $p^{\mathrm{jet}}_{\mathrm{T,lead}}$ *(b),* $m^{\gamma j}$ *(c) and* $|\cos\theta^{\gamma j}|$ *(d) for the photon+one-jet sample.*

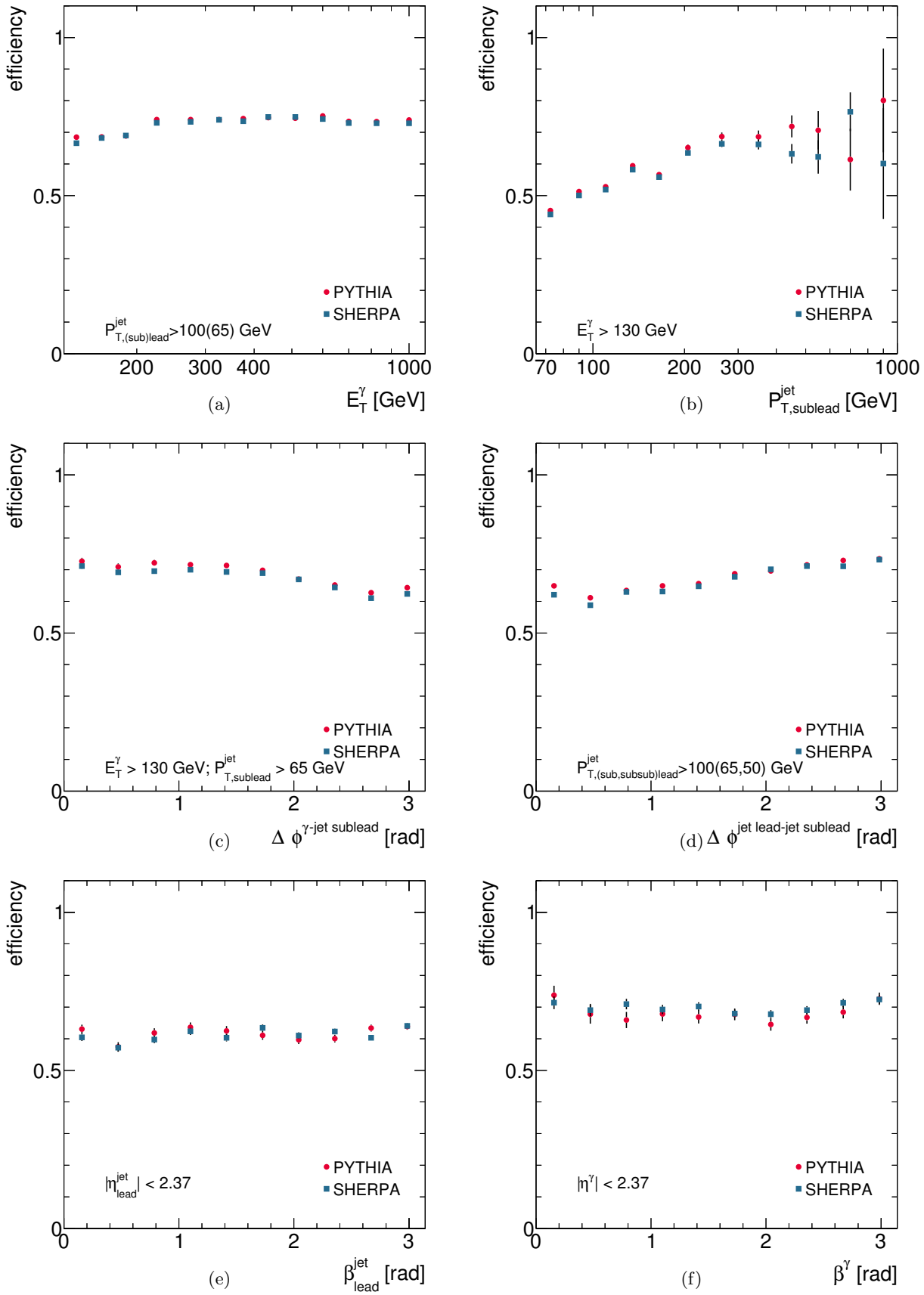


Figure 6.48: *Selection efficiency from* PYTHIA *(dots) and* SHERPA *(squares) as a function of* E_{T}^{γ} *(a),* $p_{\mathrm{T,sublead}}^{\mathrm{jet}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$ *(d),* $\beta_{\mathrm{lead}}^{\mathrm{jet}}$ *(e) and* β^{γ} *(f) for the photon+two-jets (a,b,c,d) and the colour coherence (e,f) samples.*

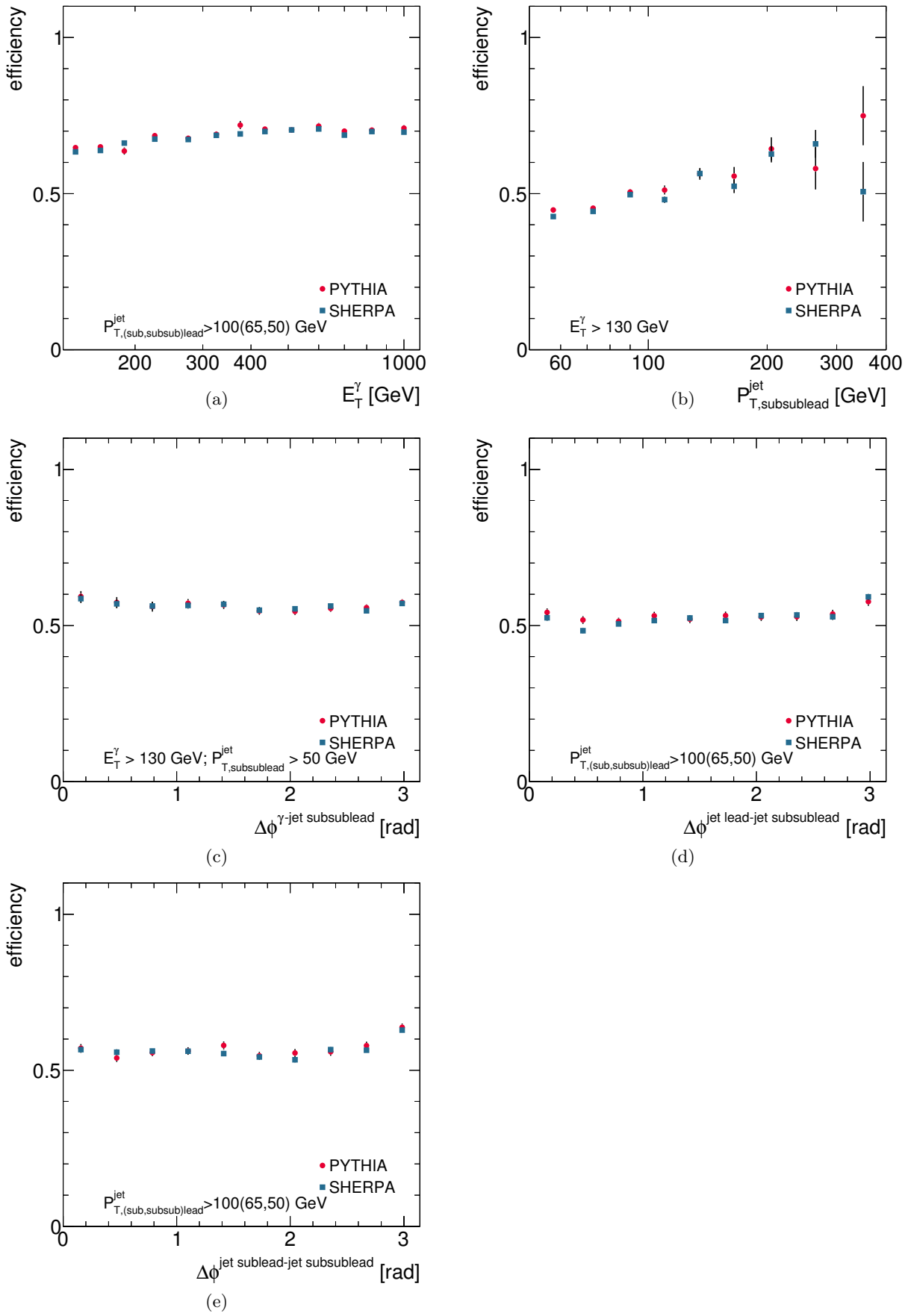


Figure 6.49: *Selection efficiency from* PYTHIA *(dots) and* SHERPA *(squares) as a function of* E^{γ}_{T} *(a),* $p^{\mathrm{jet}}_{\mathrm{T,subsublead}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\,subsublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,subsublead}}$ *(d),* $\Delta\phi^{\mathrm{jet\,sublead}-\mathrm{jet\,subsublead}}$ *(e) for the photon+three-jets sample.*

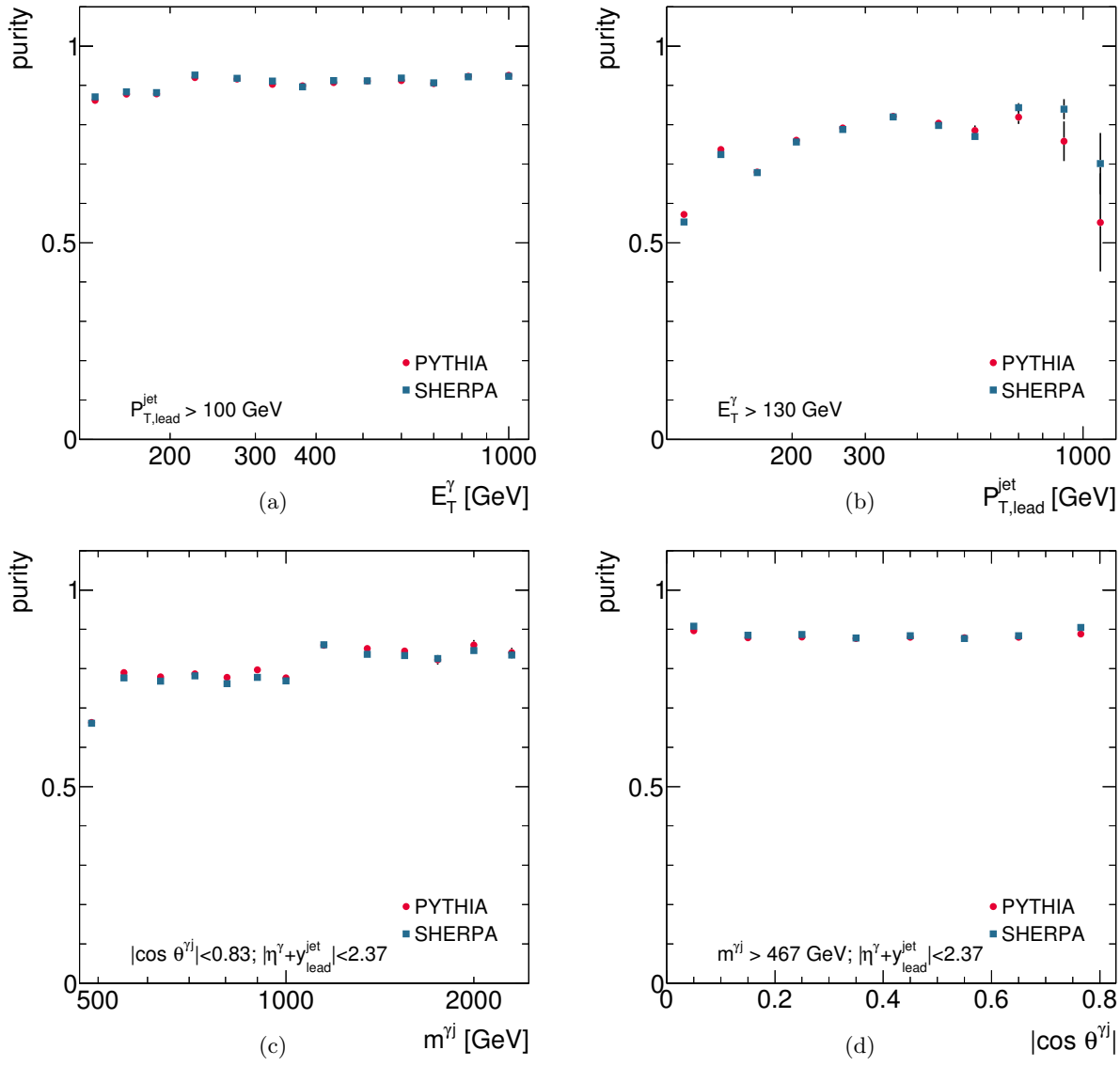


Figure 6.50: *Selection purity from* PYTHIA *(dots) and* SHERPA *(squares) as a function of* E_{T}^{γ} *(a),* $p_{\mathrm{T,lead}}^{\mathrm{jet}}$ *(b),* $m^{\gamma j}$ *(c) and* $|\cos\theta^{\gamma j}|$ *(d) for the photon+one-jet sample.*

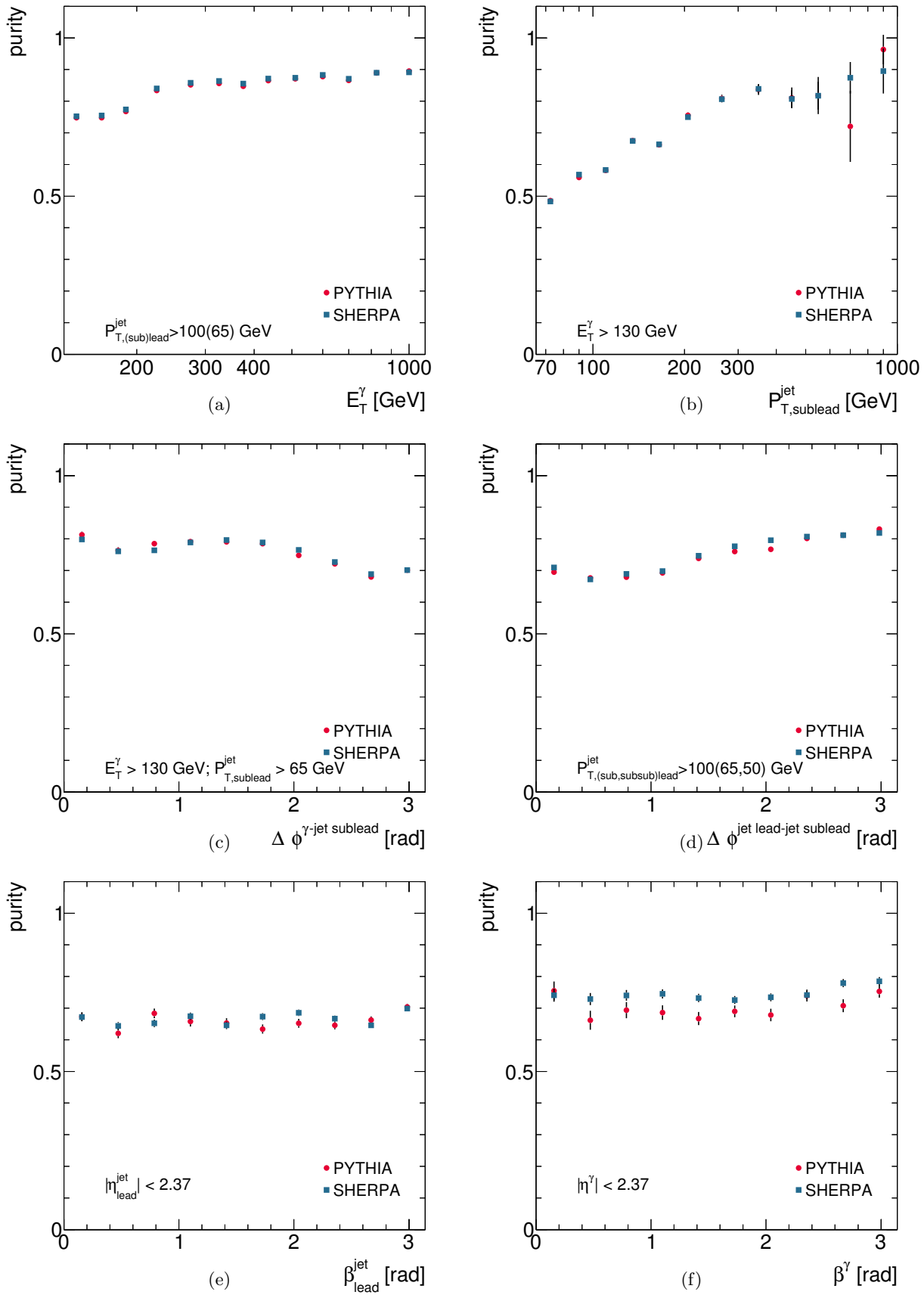


Figure 6.51: *Selection purity from* PYTHIA *(dots) and* SHERPA *(squares) as a function of* E_{T}^{γ} *(a),* $p_{\mathrm{T,sublead}}^{\mathrm{jet}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$ *(d),* $\beta_{\mathrm{lead}}^{\mathrm{jet}}$ *(e) and* β^{γ} *(f) for the photon+two-jets (a,b,c,d) and the colour coherence (e,f) samples.*

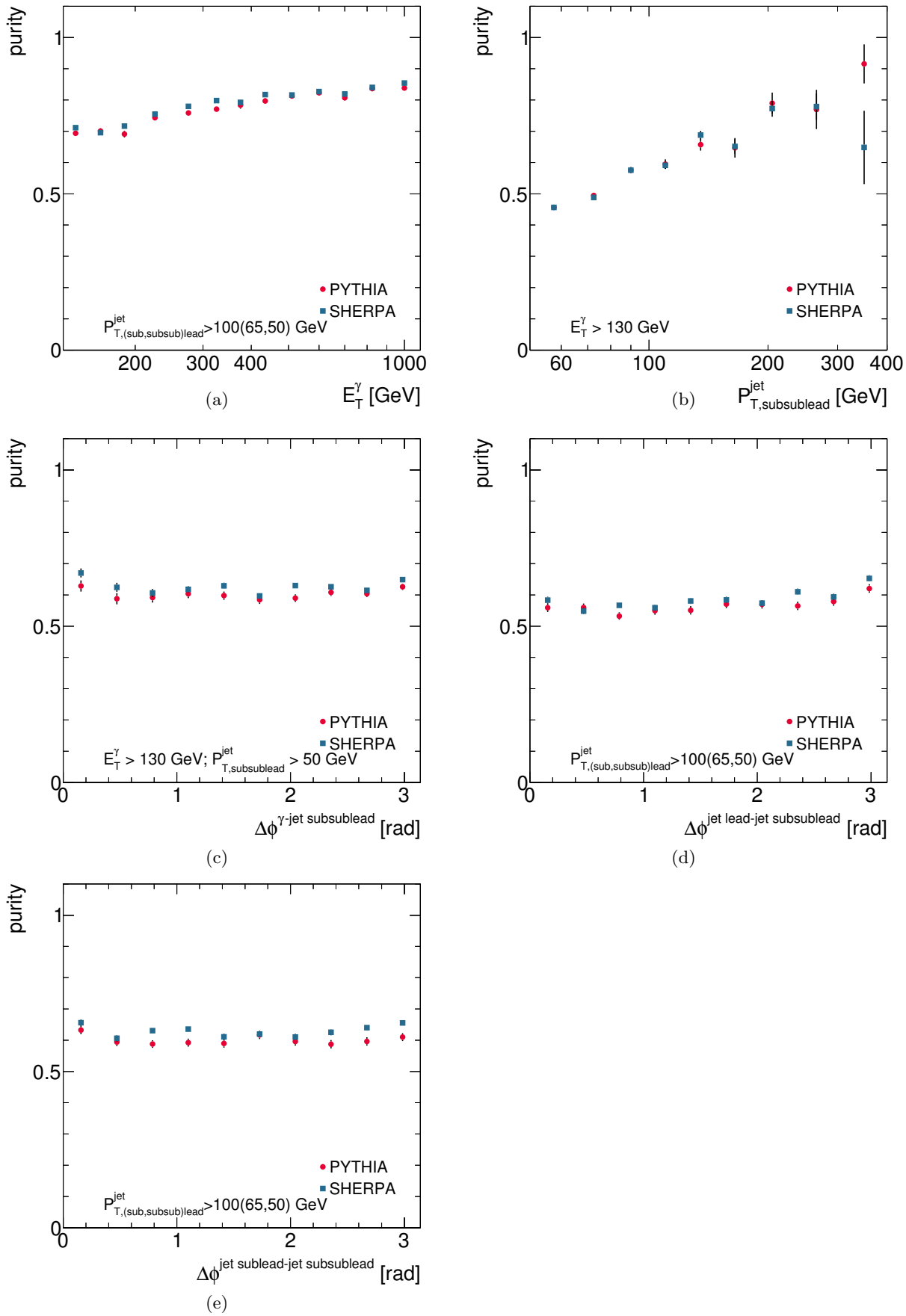


Figure 6.52: *Selection purity from* PYTHIA *(dots) and* SHERPA *(squares) as a function of* E^{γ}_{T} *(a),* $p^{\mathrm{jet}}_{\mathrm{T,subsublead}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\ subsublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\ lead}-\mathrm{jet\ subsublead}}$ *(d),* $\Delta\phi^{\mathrm{jet\ sublead}-\mathrm{jet\ subsublead}}$ *(e) for the photon+three-jets sample.*

6.10 Cross section measurements

Isolated-photon plus jets cross sections were measured for the phase-space defined by the particle level selection defined in section 6.3.1. For each sample, the phase-space was defined by the requirements defined in section 6.4 applied over particle level objects. The cross section was measured as a function of the following observables:

- **photon+one-jet sample:** E_{T}^{γ} and $p_{\mathrm{T,lead}}^{\mathrm{jet}}$;
- **photon+one-jet $m^{\gamma j}$-cos $\theta^{\gamma j}$ sample:** $m^{\gamma j}$ and $|\cos\theta^{\gamma j}|$;
- **photon+two-jets sample:** E_{T}^{γ}, $p_{\mathrm{T,sublead}}^{\mathrm{jet}}$, $\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$ and $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$;
- **colour coherence samples:** β^{γ} and $\beta_{\mathrm{lead}}^{\mathrm{jet}}$;
- **photon+three-jets sample:** E_{T}^{γ}, $p_{\mathrm{T,subsublead}}^{\mathrm{jet}}$, $\Delta\phi^{\gamma-\mathrm{jet\,subsublead}}$, $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,subsublead}}$ and $\Delta\phi^{\mathrm{jet\,sublead}-\mathrm{jet\,subsublead}}$.

The data distributions, after background subtraction, were corrected to the particle level using bin-by-bin acceptance correction factors determined using the MC samples. These corrections factors take into account the efficiency of the selection criteria and the purity and efficiency of the jet reconstruction and the efficiency of the photon reconstruction. For this approach to be valid, the uncorrected distributions of the data must be adequately described by the MC simulations at the reconstruction level. This condition was generally satisfied by both the Pythia and Sherpa MC samples except for the β^{γ} and β^{γ} distributions, for which the description of the data distributions by Pythia is rather poor.

The data distributions were corrected to the particle level using

$$\frac{d\sigma}{dA}(i) = \frac{N_A^{\mathrm{sig}}(i)\ C^{\mathrm{MC}}(i)}{\mathcal{L}\ \Delta A(i)}, \tag{6.10}$$

where $(d\sigma/dA)(i)$ is the differential cross section as a function of observable A, $N_A^{\mathrm{sig}}(i)$ is the number of background-subtracted data events in bin i, $C^{\mathrm{MC}}(i)$ is the correction factor in bin i, $\mathcal{L}$ is the integrated luminosity and $\Delta A(i)$ is the width of bin i. The acceptance correction factors were computed from Pythia as

$$C^{\mathrm{MC}}(i) = \frac{\alpha\ N_{\mathrm{part}}^{\mathrm{MC,H}}(i) + (1-\alpha)\ N_{\mathrm{part}}^{\mathrm{MC,B}}(i)}{\alpha\ N_{\mathrm{reco}}^{\mathrm{MC,H}}(i) + (1-\alpha)\ N_{\mathrm{reco}}^{\mathrm{MC,B}}(i)}, \tag{6.11}$$

and Sherpa as

$$C^{\mathrm{MC}}(i) = \frac{N_{\mathrm{part}}^{\mathrm{MC}}(i)}{N_{\mathrm{reco}}^{\mathrm{MC}}(i)}, \tag{6.12}$$

where α is the value obtained from the fit to the data distribution of each observable (see table 6.1). Pythia was used as the nominal MC for the correction factors, except for the $\beta_{\mathrm{lead}}^{\mathrm{jet}}$ and β^{γ} distributions, for which Sherpa was used. To be consistent, $N_A^{\mathrm{sig}}(i)$ was obtained using the leakage fractions from the same MC used for the correction factors.

The deviations in the results obtained by using Sherpa (or Pythia in the case of the $\beta_{\mathrm{lead}}^{\mathrm{jet}}$ and β^{γ} cross sections) to correct the data were taken to represent systematic uncertainties of the effect of the QCD-cascade and hadronisation models in the corrections (see section 6.11). The acceptance correction factors are shown in figures. 6.53 to 6.55. The correction factors differ from unity by typically $\approx 20\%$. The correction factors are very similar for Pythia and Sherpa.

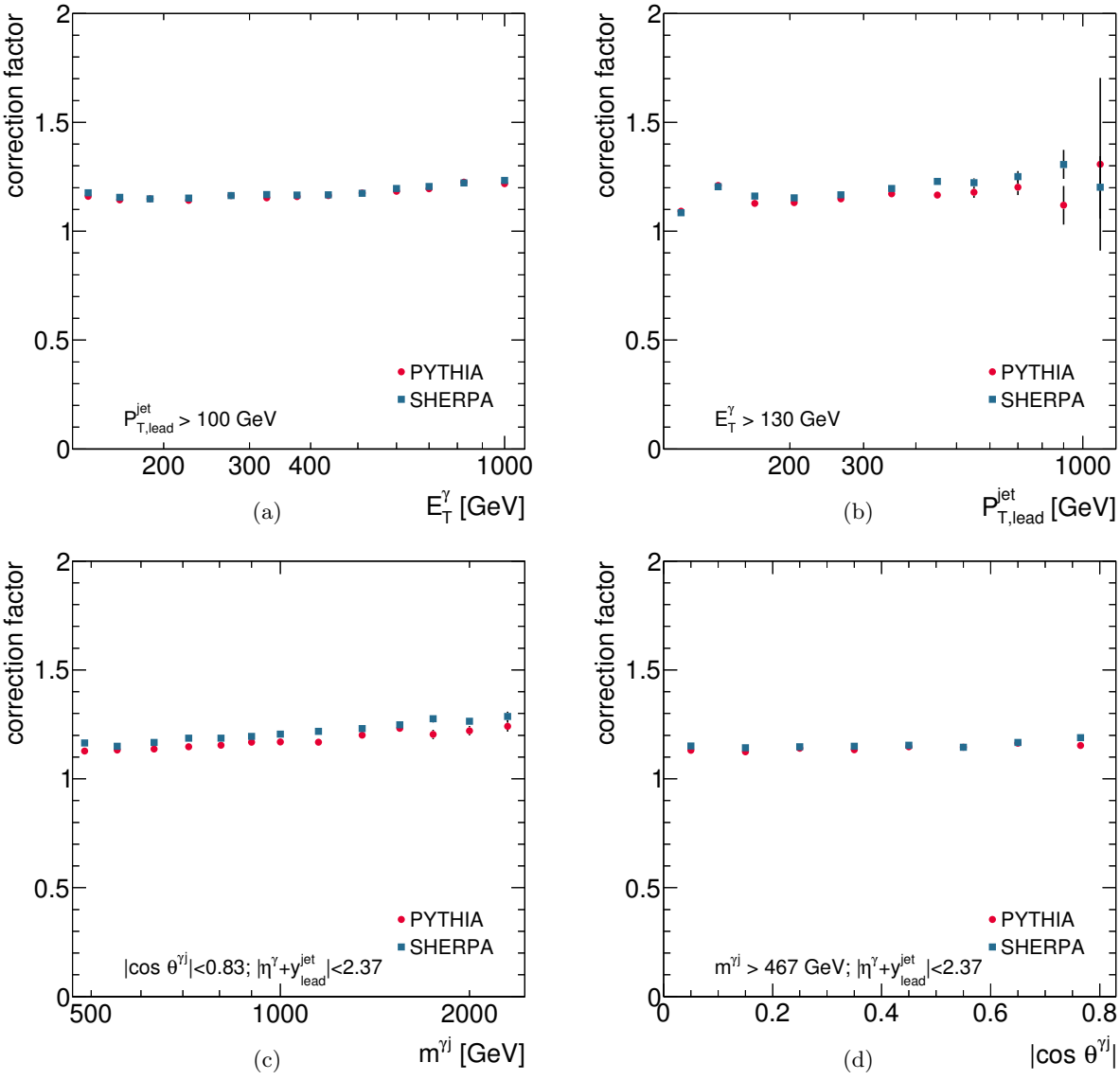


Figure 6.53: *Correction factors from* PYTHIA *(dots) and* SHERPA *(squares) as functions of* E^{γ}_{T} *(a),* $p^{\mathrm{jet}}_{\mathrm{T,lead}}$ *(b),* $m^{\gamma j}$ *(c) and* $|\cos\theta^{\gamma j}|$ *(d) for the photon+one-jet sample.*

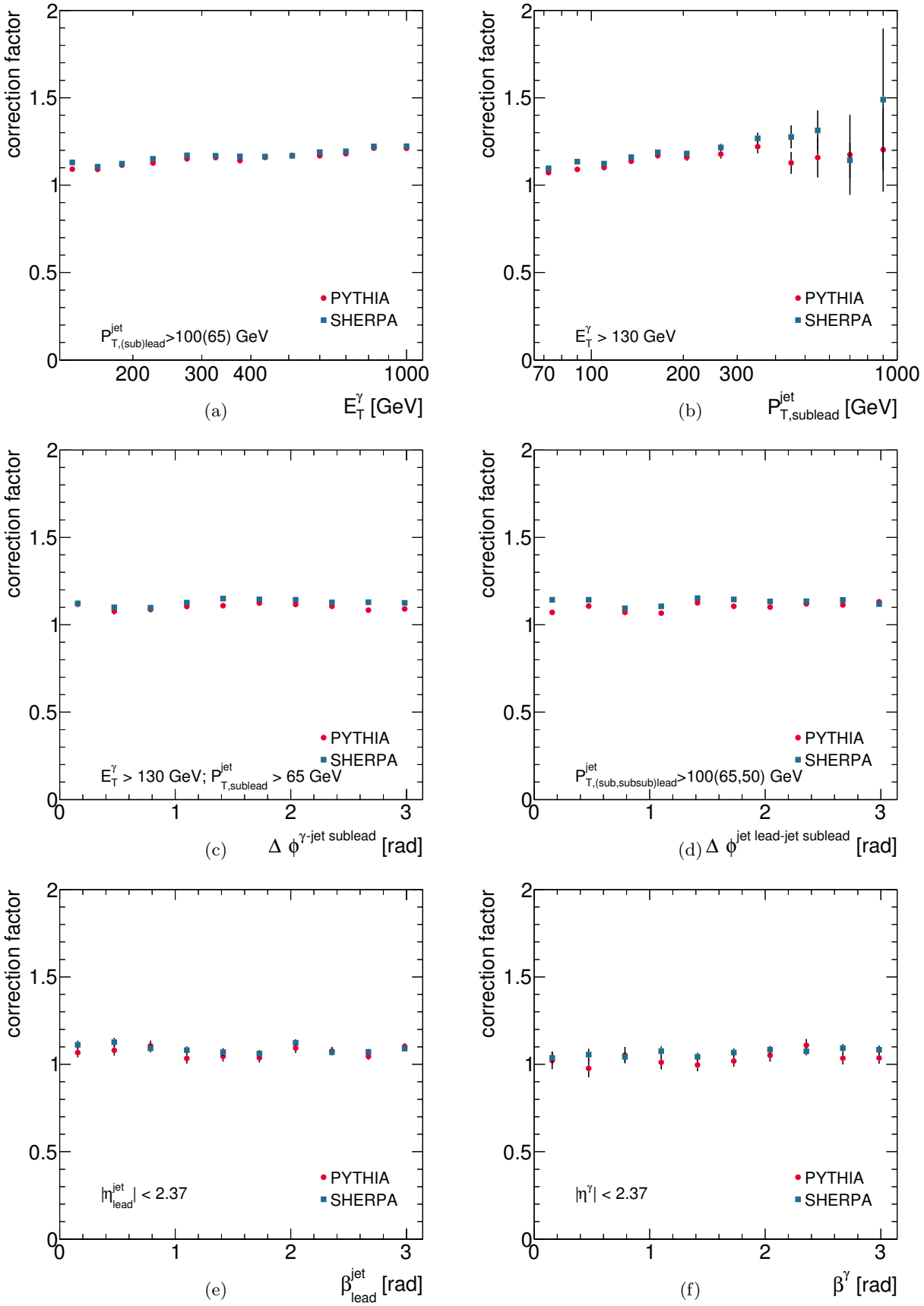


Figure 6.54: *Correction factors from* PYTHIA *(dots) and* SHERPA *(squares) as functions of* E^{γ}_{T} *(a),* $p^{\mathrm{jet}}_{\mathrm{T,sublead}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$ *(d),* $\beta^{\mathrm{jet}}_{\mathrm{lead}}$ *(e) and* β^{γ} *(f) for the photon+two-jets (a,b,c,d) and the colour coherence (e,f) samples.*

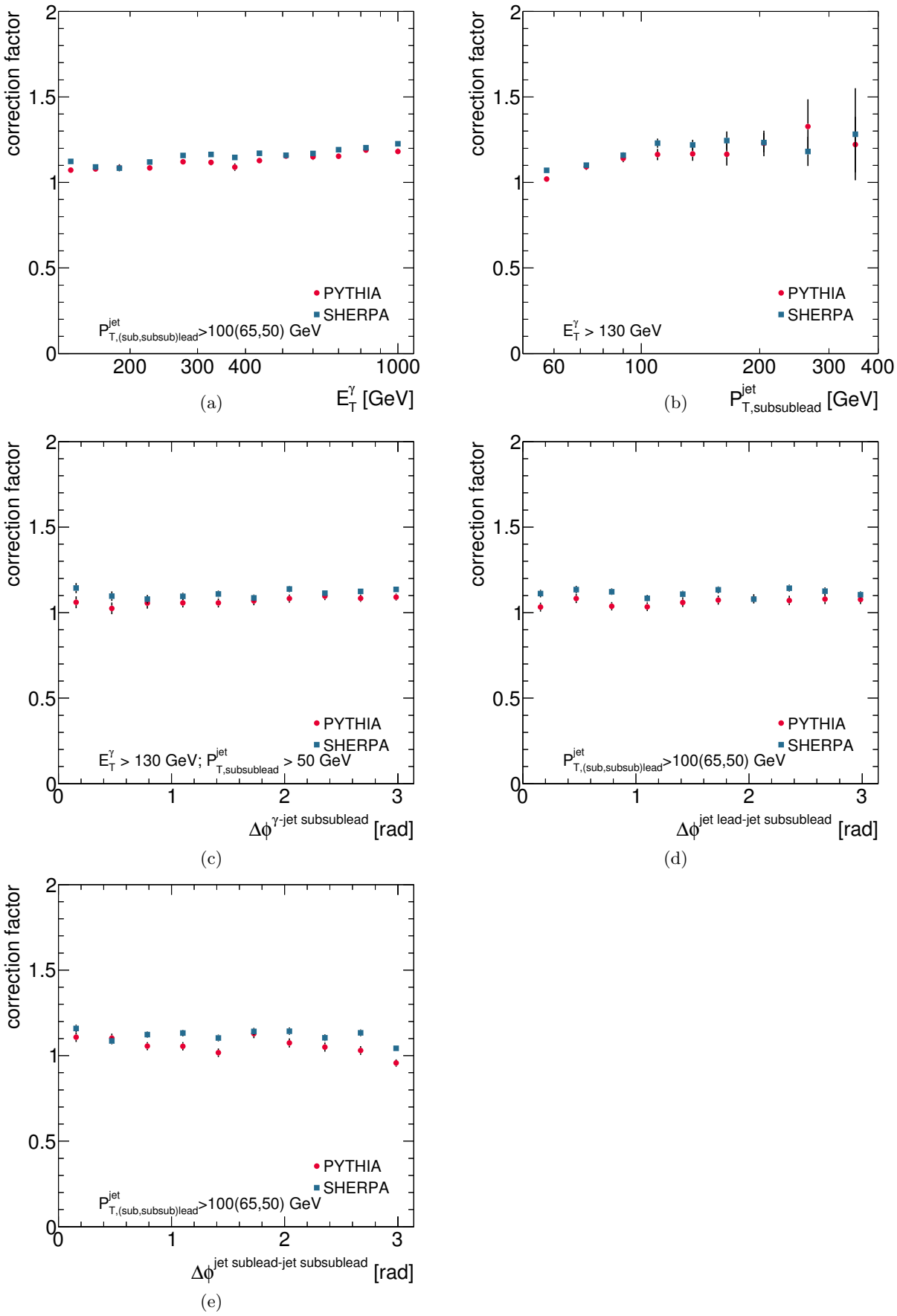


Figure 6.55: *Correction factors from* PYTHIA *(dots) and* SHERPA *(squares) as functions of* E^{γ}_{T} *(a),* $p^{\mathrm{jet}}_{\mathrm{T,subsublead}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\ subsublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\ lead}-\mathrm{jet\ subsublead}}$ *(d),* $\Delta\phi^{\mathrm{jet\ sublead}-\mathrm{jet\ subsublead}}$ *(e) for the photon+three-jets sample.*

6.11 Systematic uncertainties

Several sources of systematic uncertainties that affect the measurements were investigated. These sources include the photon energy scale and resolution, the jet energy scale and resolution, the model dependence, the photon identification efficiency, the choice of background control regions, the signal modelling and the identification and isolation correlation in the background. Each source is discussed in detail in this section. Additional cross-checks are included in chapter 7.

6.11.1 Uncertainty on the measurement of the integrated luminosity

The uncertainty on the integrated luminosity is $\pm 2.8\%$ [88]. This uncertainty is fully correlated in all bins of all the measured cross sections and it was not added in quadrature to the other uncertainties.

6.11.2 Photon energy scale and resolution

Differences between the energy scale and resolution in data and simulations lead to systematic uncertainties. A total of 20 individual components influencing the energy measurement of the photon were studied and varied within their uncertainties, up and down, to assess the overall uncertainty on the energy measurement. These parameters were propagated through the analysis separately to maintain the full information on the correlations. These components are:

- uncertainty related to the electromagnetic scale extraction from $Z \to ee$ data events, both from statistics and systematics;
- uncertainty related to the PS energy scale calibration;
- uncertainty related to the energy scale between the first and second calorimeter layers due to the fit procedure for muons and electron to muon differences;
- LAr related uncertainties from the knowledge of the internal calorimeter geometry and the effect of cross-talk between layers:
  - LAr E1/E2 relative calibration from muons;
  - LAr E1/E2 modelling for unconverted photons;
  - LAr E1/E2 modelling differences between electrons and unconverted photons;
  - LAr E1/E2 modelling for electrons for $|\eta| > 1.8$ where only electrons are used to estimate the material before the calorimeter.
- material uncertainties:
  - uncertainty on the material in the inner detector;
  - uncertainty on the cryostat passive material;
  - uncertainty on the calorimeter passive material.
- uncertainty on the gain switch correction in L1 and L2;
- uncertainty on E1/E2 from the different shower shapes description in various physics modelling options in Geant4;
- mismodelling of the pedestal;

- uncertainty on the lateral leakage MC description for converted and unconverted photons;
- uncertainty from true converted/unconverted photons seen as unconverted/converted respectively;
- uncertainty from conversion radius mismodeling.

Similarly to the energy scale uncertainty, the energy resolution is also influenced by different contributions (seven components), which were also propagated through the analysis separately to maintain the full information on the correlations. These components are:

- mismodelling of the resolution sampling term, the electronics noise term, the asymptotic resolution at high energy and the effect of passive material upstream of the calorimeter;
- material uncertainty for the inner detector, the calorimeter, the cryostat and the gap between the barrel and end cap parts of the ATLAS detector;
- uncertainty based on different pile-up conditions.

The relative uncertainties in the photon energy scale and due to the resolution are shown in Fig. 6.56 as functions of E_{T}^{γ}. The systematic uncertainties on the measured cross sections due to the effects mentioned above were estimated by varying each individual source of uncertainty separately in the MC simulations and then added in quadrature. Figures 6.57 to 6.62 show the resulting uncertainties as functions of the observables studied.

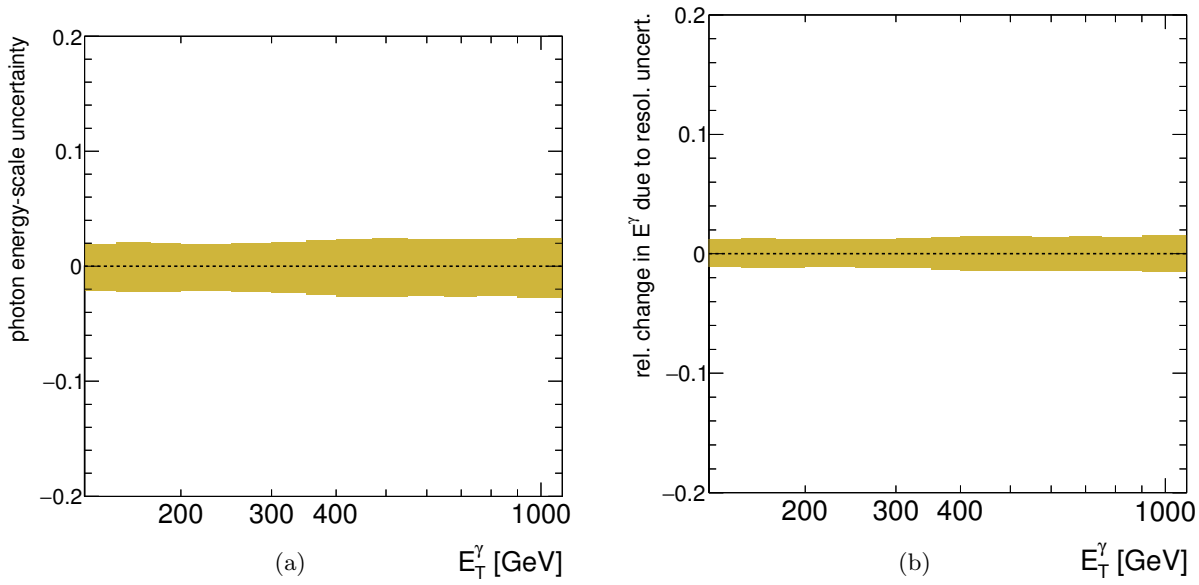


Figure 6.56: *Relative photon energy-scale uncertainty (a) and relative change in photon energy due to the resolution uncertainty as functions of* E_{T}^{γ}.

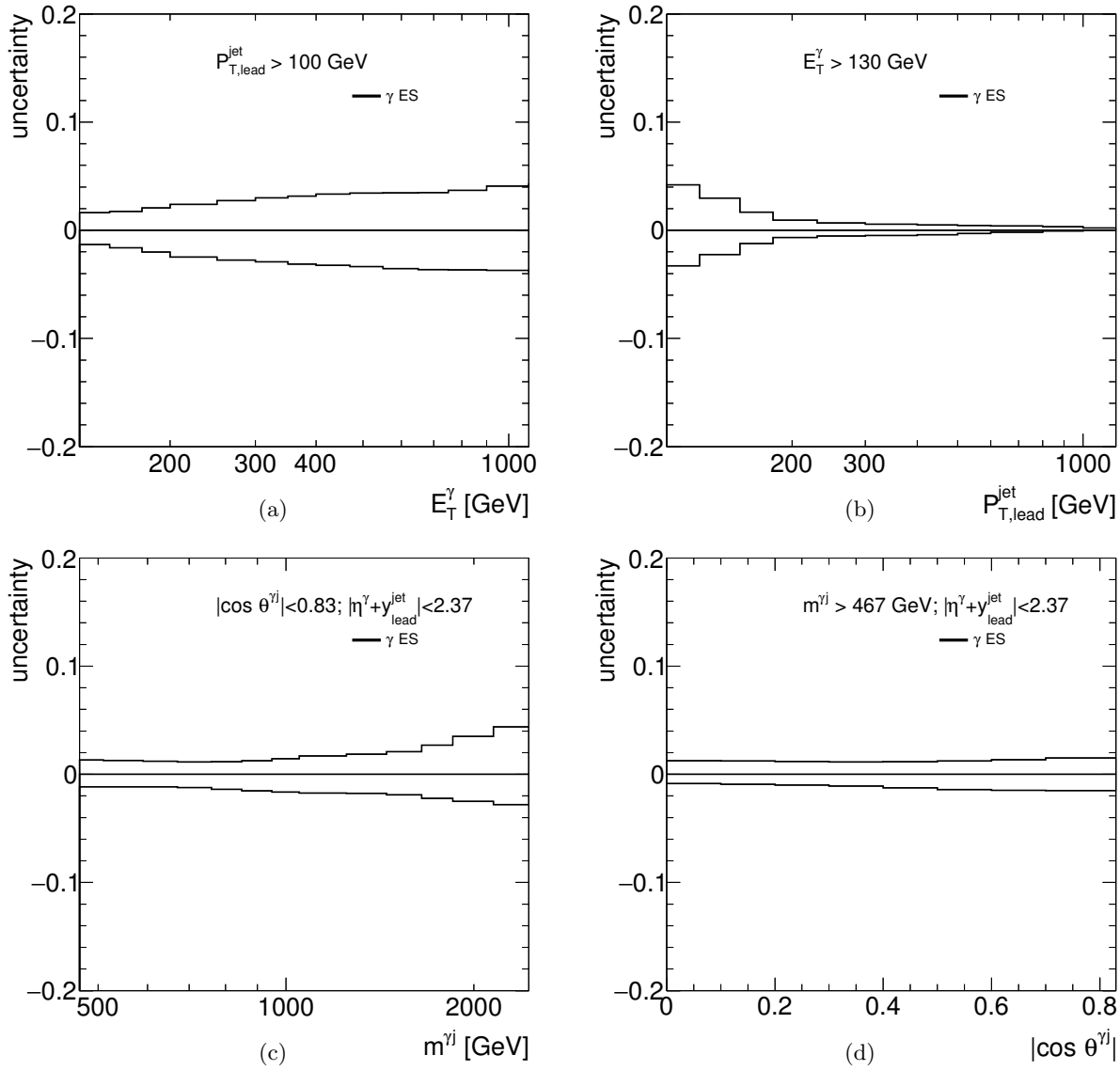


Figure 6.57: *Systematic uncertainties on the measured cross section due to the uncertainty in the photon energy scale as functions of* E^{γ}_{T} *(a),* $p^{\mathrm{jet}}_{\mathrm{T,lead}}$ *(b),* $m^{\gamma j}$ *(c) and* $|\cos\theta^{\gamma j}|$ *(d) for the photon+one-jet sample.*

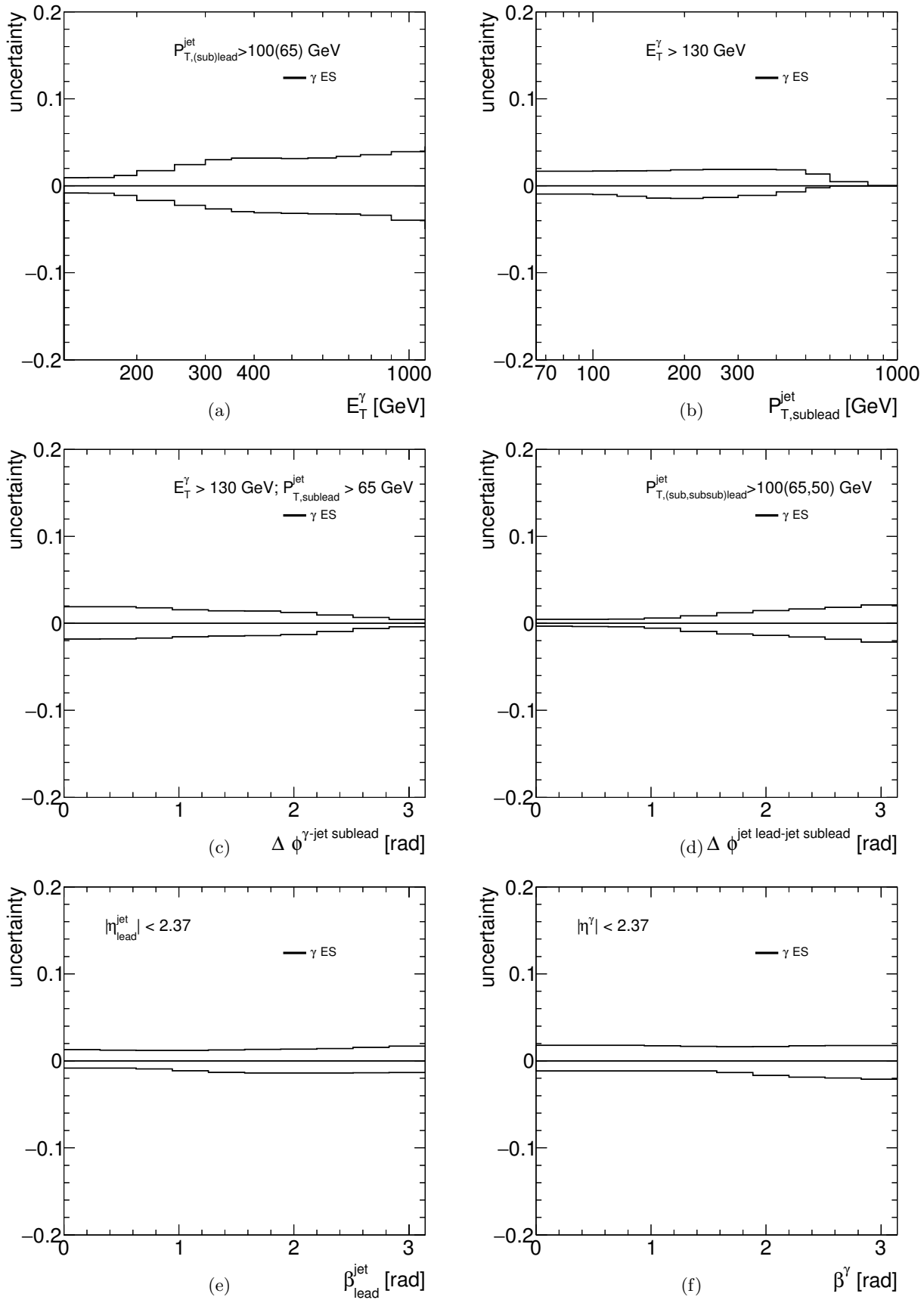


Figure 6.58: *Systematic uncertainties on the measured cross section due to the uncertainty in the photon energy scale as functions of* E^{γ}_{T} *(a),* $p^{\mathrm{jet}}_{\mathrm{T,sublead}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$ *(d),* $\beta^{\mathrm{jet}}_{\mathrm{lead}}$ *(e) and* β^{γ} *(f) for the photon+two-jets (a,b,c,d) and the colour coherence (e,f) samples.*

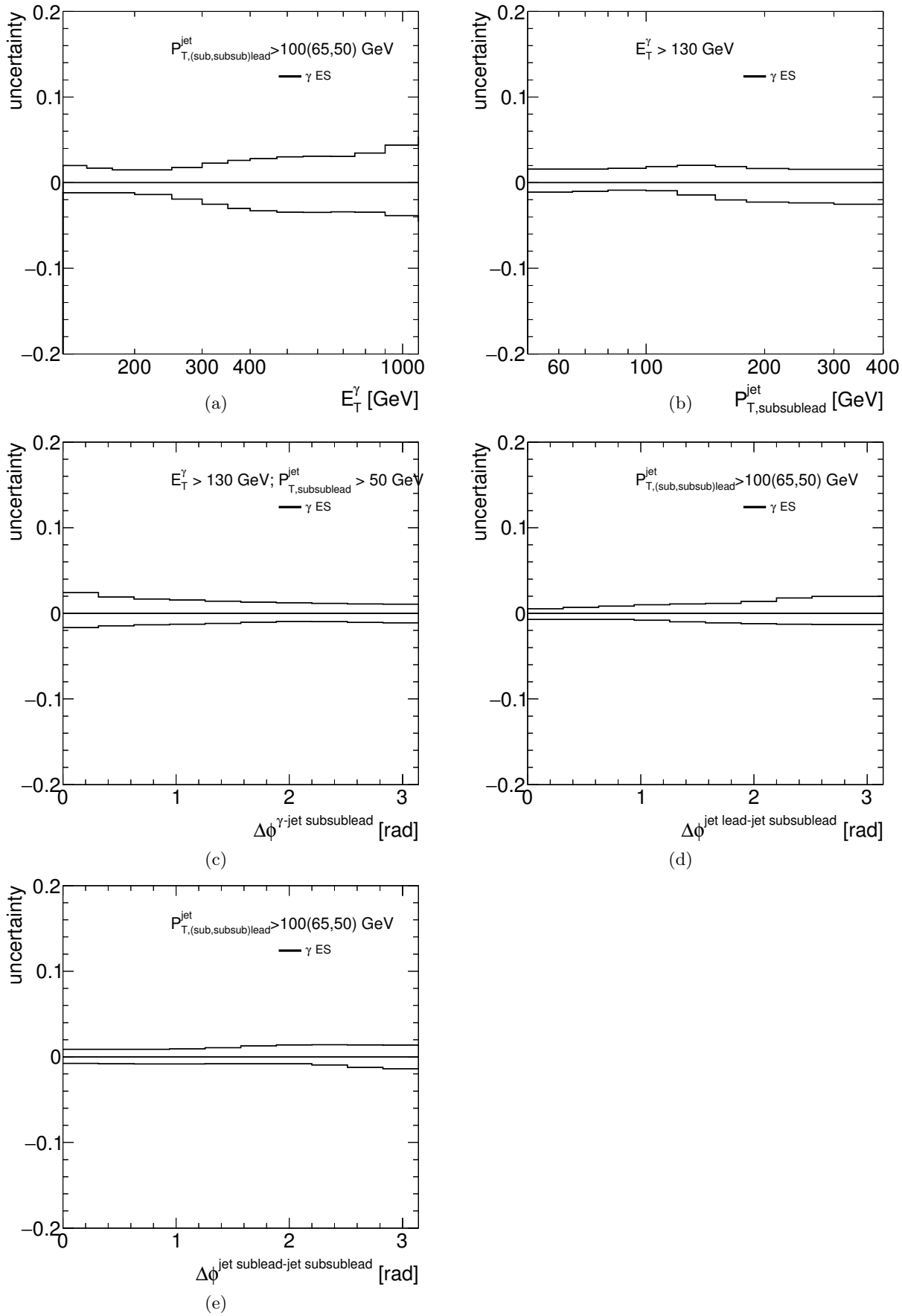


Figure 6.59: *Systematic uncertainties on the measured cross section due to the uncertainty in the photon energy scale as functions of* E_{T}^{γ} *(a),* $p_{\mathrm{T,subsublead}}^{\mathrm{jet}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\ subsublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\ lead}-\mathrm{jet\ subsublead}}$ *(d),* $\Delta\phi^{\mathrm{jet\ sublead}-\mathrm{jet\ subsublead}}$ *(e) for the photon+three-jets sample.*

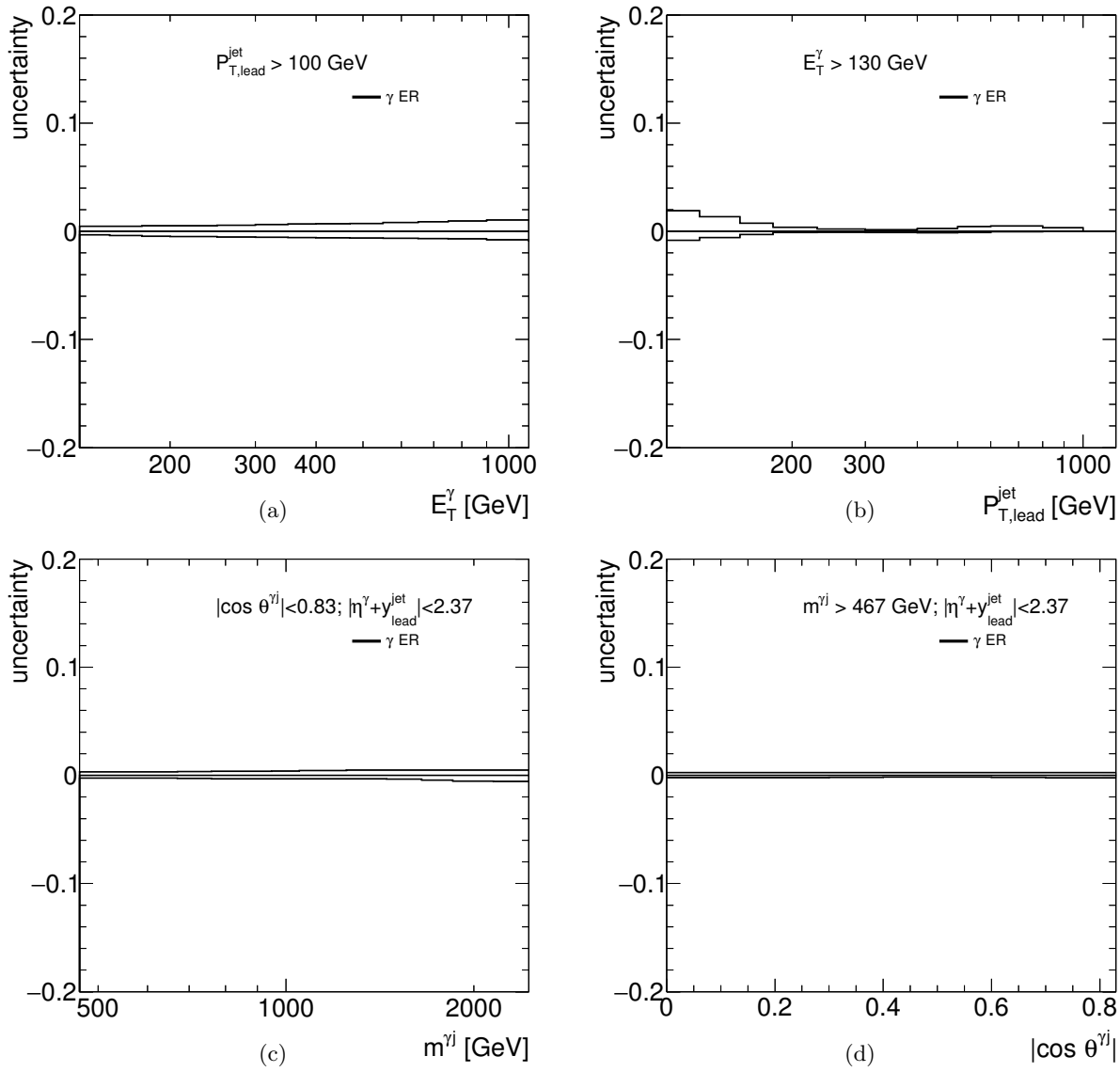


Figure 6.60: *Systematic uncertainties on the measured cross section due to the uncertainty in the photon energy resolution as functions of* E_{T}^{γ} *(a),* $p_{\mathrm{T,lead}}^{\mathrm{jet}}$ *(b),* $m^{\gamma j}$ *(c) and* $|\cos\theta^{\gamma j}|$ *(d) for the photon+one-jet sample.*

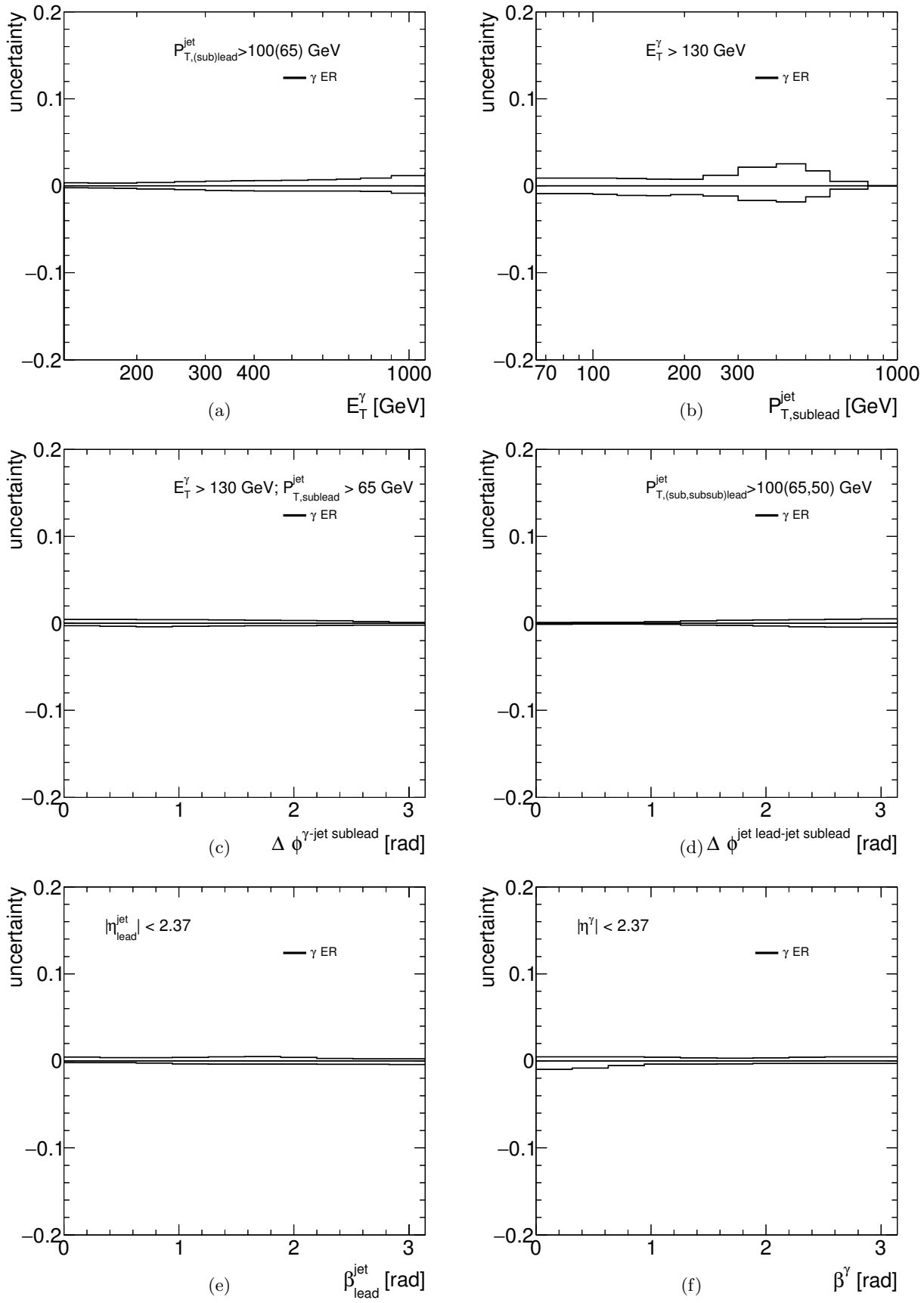


Figure 6.61: *Systematic uncertainties on the measured cross section due to the uncertainty in the photon energy resolution as functions of* E^{γ}_{T} *(a),* $p^{\mathrm{jet}}_{\mathrm{T,sublead}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$ *(d),* $\beta^{\mathrm{jet}}_{\mathrm{lead}}$ *(e) and* β^{γ} *(f) for the photon+two-jets (a,b,c,d) and the colour coherence (e,f) samples.*

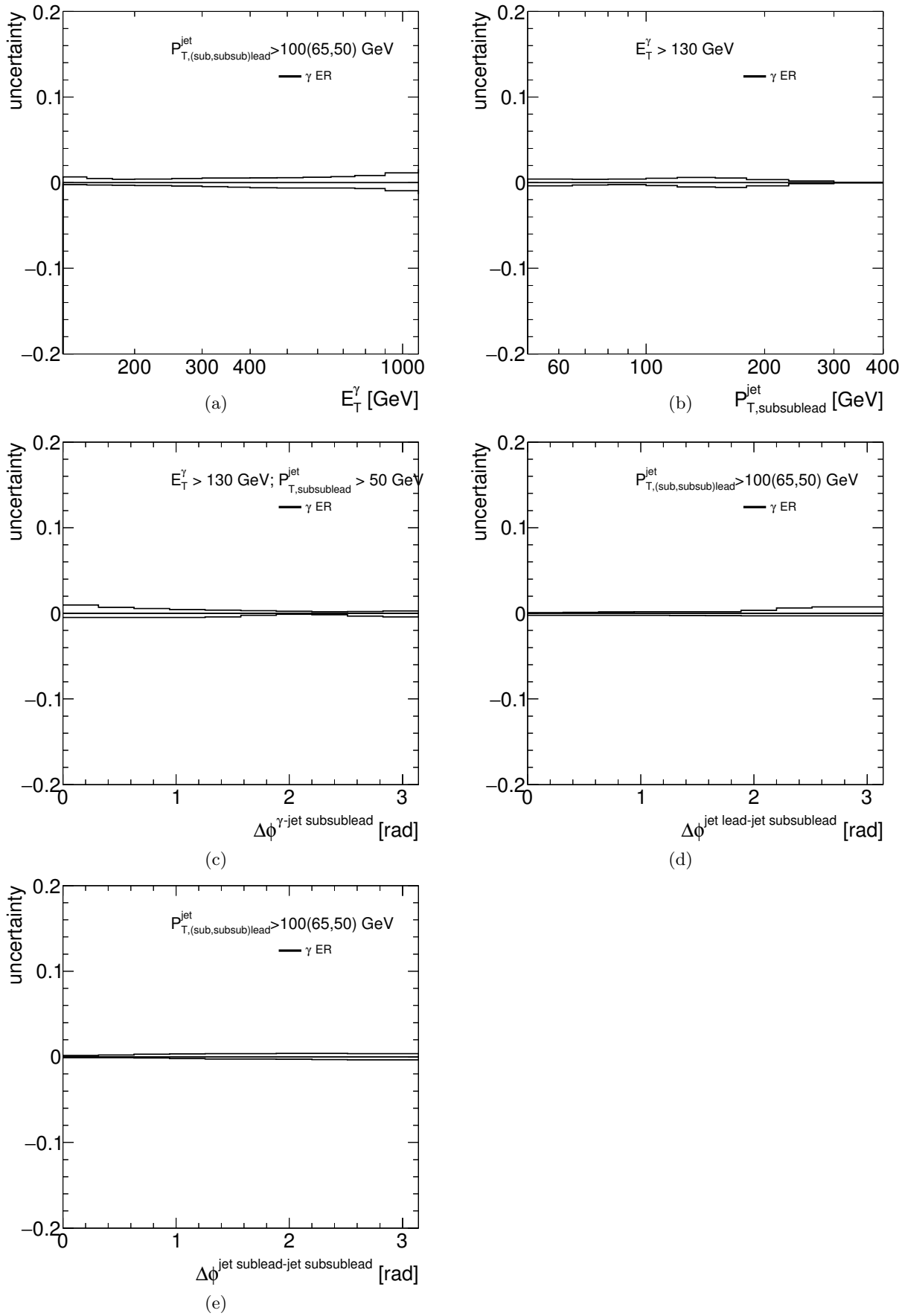


Figure 6.62: *Systematic uncertainties on the measured cross section due to the uncertainty in the photon energy resolution as functions of* E_{T}^{γ} *(a),* $p_{\mathrm{T,subsublead}}^{\mathrm{jet}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\,subsublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,subsublead}}$ *(d),* $\Delta\phi^{\mathrm{jet\,sublead}-\mathrm{jet\,subsublead}}$ *(e) for the photon+three-jets sample.*

6.11.3 Jet energy scale

Differences between the energy scale in data and simulations lead to systematic uncertainties. The final iteration of the JES uncertainty for the full 2012 dataset contains a full treatment of bin-to-bin correlations for systematic uncertainties. This was achieved through the splitting of the nuisance parameters coming from the various in-situ techniques. A total of 67 individual components influencing the energy measurement of the jets were studied and varied within their uncertainties to assess the overall uncertainty on the jet energy measurement. These components were propagated through the analysis separately to maintain the full information on the correlations. These components are:

- 56 nuisance parameters from the in-situ analyses ($Z + jet$, $\gamma + jet$ and multi-jet balance);
- 2 nuisance parameters from the η intercalibration (modelling and statistics/method);
- 1 nuisance parameter from the behaviour of high-p_T jets in propagation of single hadron uncertainties to the jets;
- 1 nuisance parameter due to the differences between the sample used to derive the calibrations and the sample used in the analysis;
- 4 nuisance parameters from pile-up (3 of which are μ/NPV dependent);
- 1 nuisance parameter from the flavour composition of the sample;
- 1 nuisance parameter from the flavour response uncertainty;
- 1 nuisance parameter for punch-through jets.

The relative uncertainties in the jet energy scale are shown in figures 6.63 as functions of $p_{\mathrm{T}}^{\mathrm{jet}}$ and $|y^{\mathrm{jet}}|$. The systematic uncertainties on the measured cross sections due to the effects mentioned above were estimated by varying each individual source of uncertainty separately in the MC simulations and then added in quadrature. Figures 6.64 to 6.66 show the resulting uncertainties as functions of the observables studied.

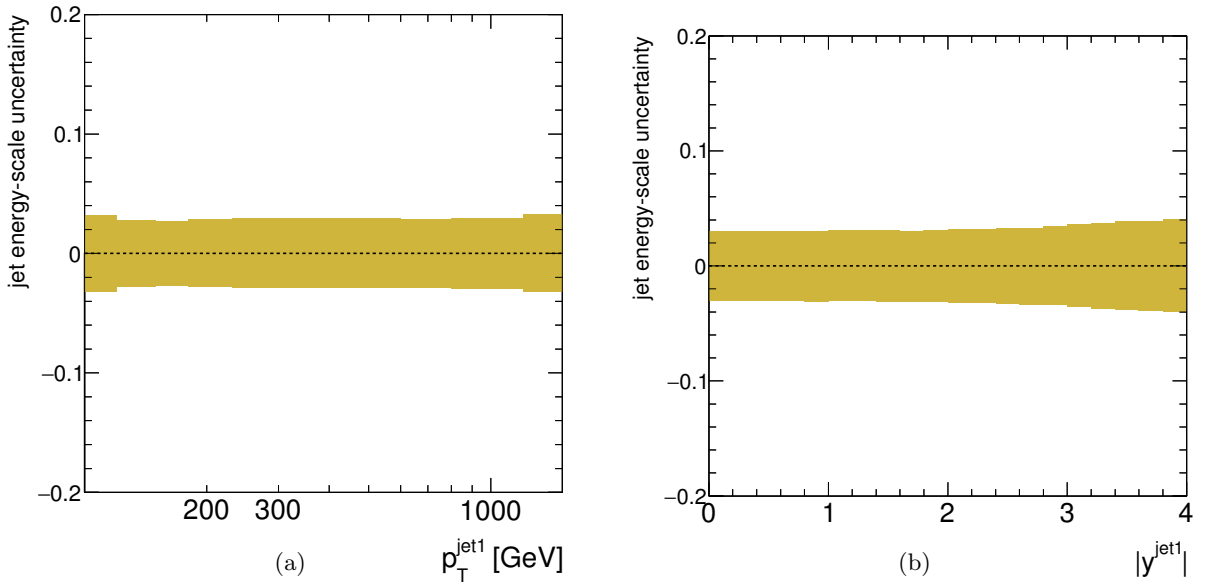


Figure 6.63: *Jet energy scale uncertainties as functions of* $p_{\mathrm{T,lead}}^{\mathrm{jet}}$ *(a) and* $|y^{\mathrm{jet}}|$ *(b).*

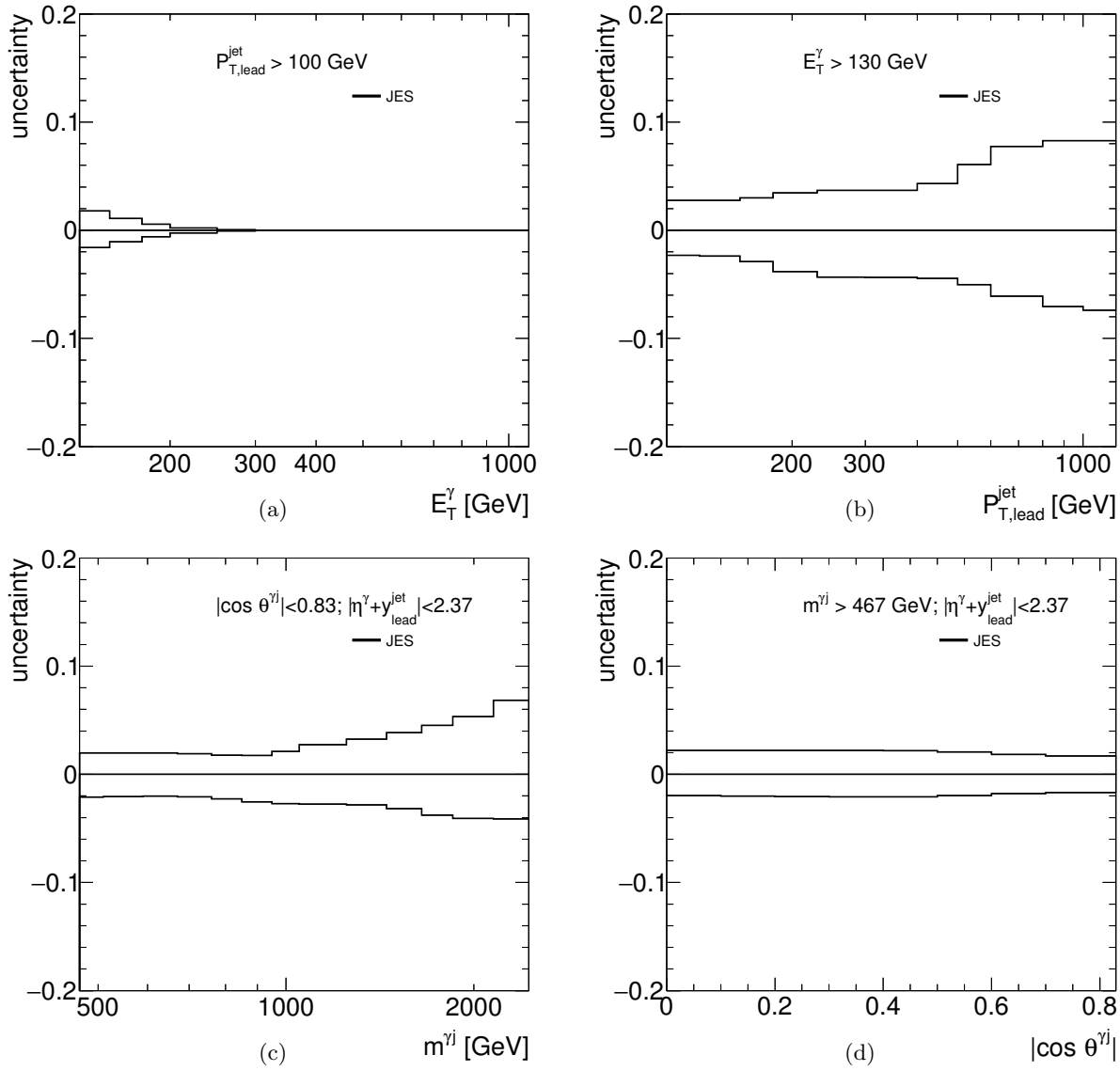


Figure 6.64: *Systematic uncertainties on the measured cross section due to the uncertainty in the jet energy scale as functions of* E_{T}^{γ} *(a),* $p_{\mathrm{T,lead}}^{\mathrm{jet}}$ *(b),* $m^{\gamma j}$ *(c) and* $|\cos\theta^{\gamma j}|$ *(d) for the photon+one-jet sample.*

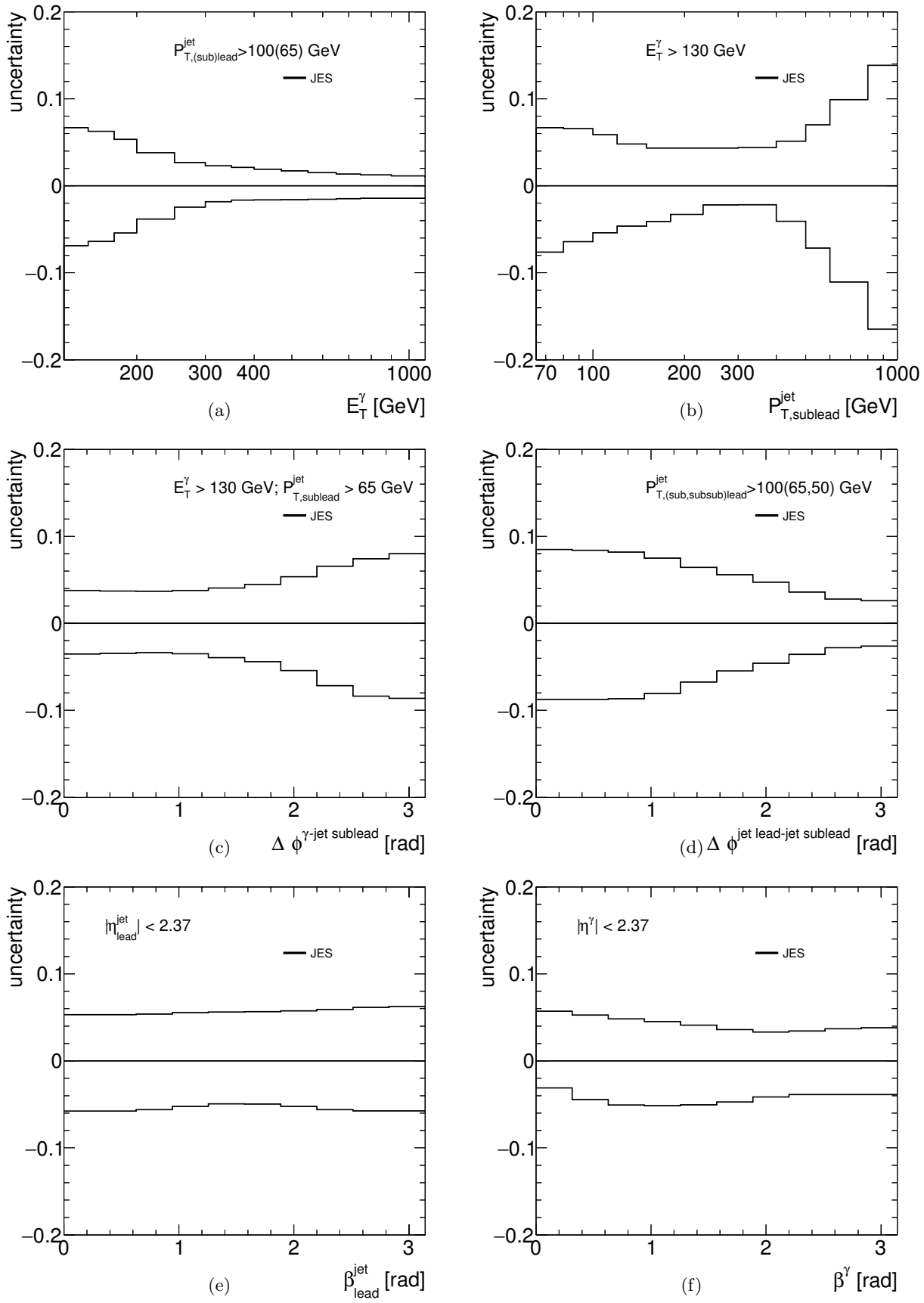


Figure 6.65: *Systematic uncertainties on the measured cross section due to the uncertainty in the jet energy scale as functions of* E^{γ}_{T} *(a),* $p^{\mathrm{jet}}_{\mathrm{T,sublead}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$ *(d),* $\beta^{\mathrm{jet}}_{\mathrm{lead}}$ *(e) and* β^{γ} *(f) for the photon+two-jets (a,b,c,d) and the colour coherence (e,f) samples.*

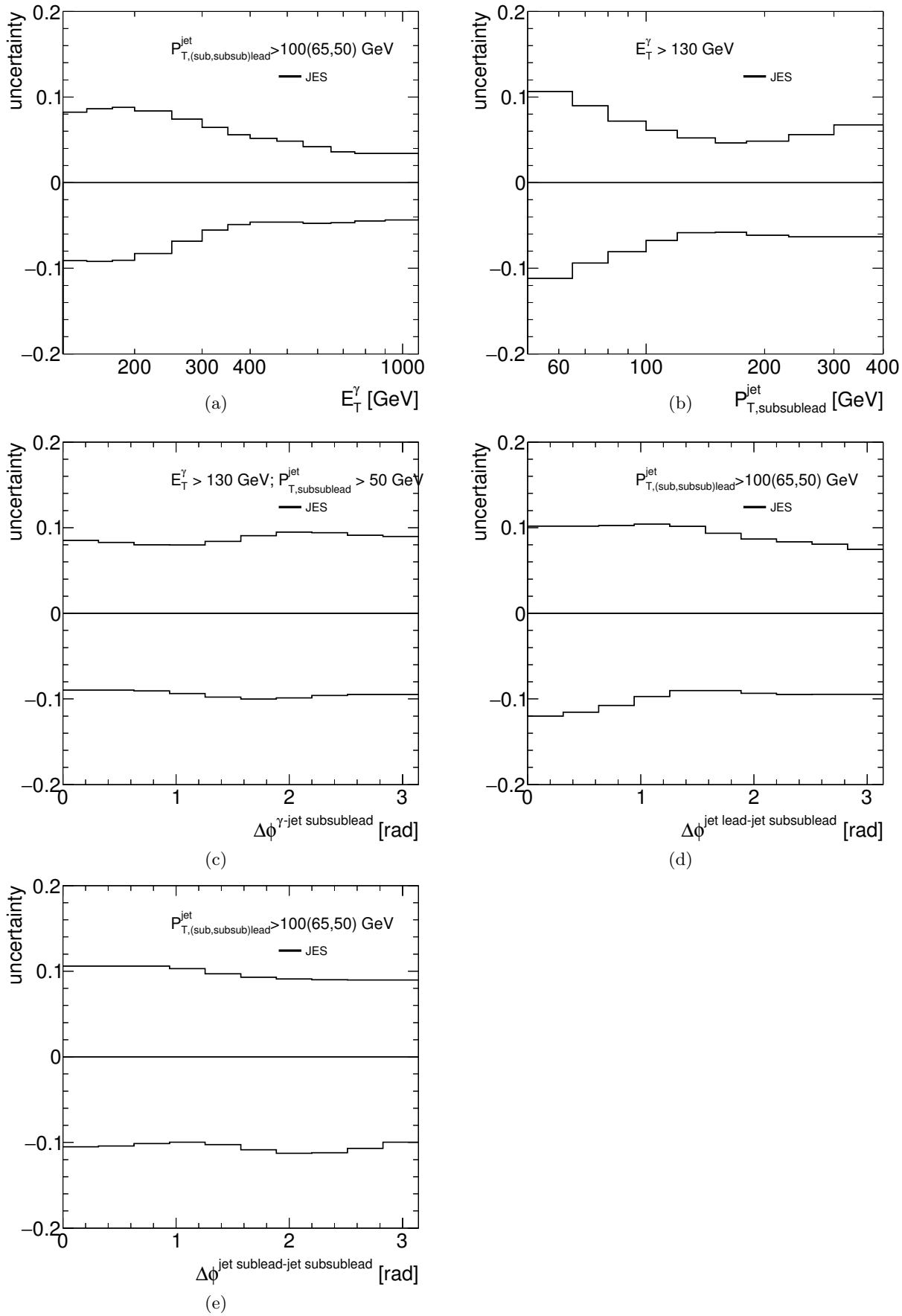


Figure 6.66: *Systematic uncertainties on the measured cross section due to the uncertainty in the jet energy scale as functions of* E_{T}^{γ} *(a),* $p_{\mathrm{T,subsublead}}^{\mathrm{jet}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\ subsublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\ lead}-\mathrm{jet\ subsublead}}$ *(d),* $\Delta\phi^{\mathrm{jet\ sublead}-\mathrm{jet\ subsublead}}$ *(e) for the photon+three-jets sample.*

6.11.4 Jet energy resolution

The Jet Energy Resolution (JER) source of systematic accounts for the fact that the simulated sample has better energy resolution than the data. The impact of this difference was estimated by smearing the jet energy in the MC simulated events and comparing the smeared and non-smeared results. Figures 6.67 to 6.69 show the resulting uncertainties as functions of the observables studied.

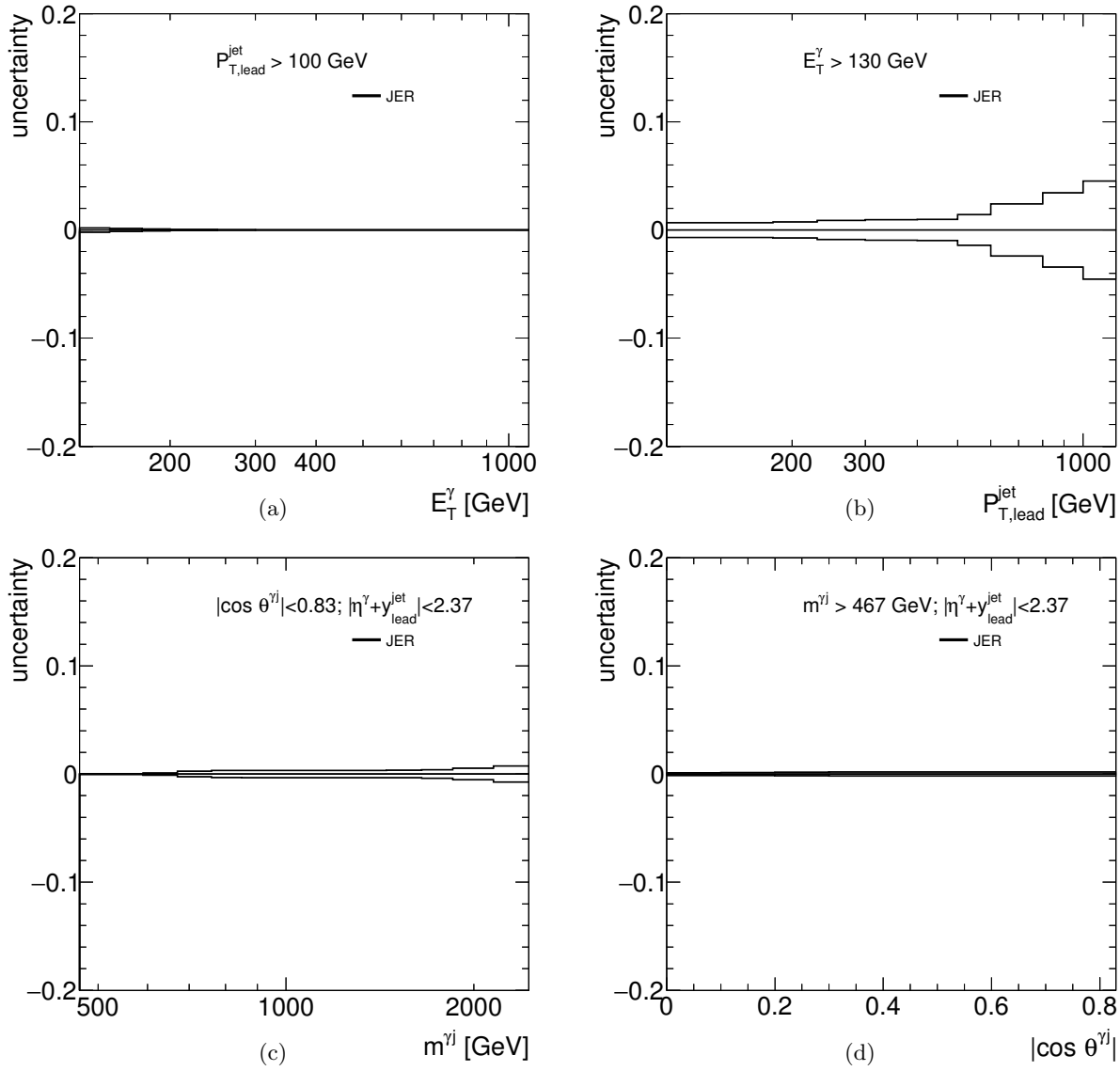


Figure 6.67: *Systematic uncertainties on the measured cross section due to the uncertainty in the jet energy resolution as functions of E^{γ}_{T} (a), $p^{\mathrm{jet}}_{\mathrm{T,lead}}$ (b), $m^{\gamma j}$ (c) and $|\cos\theta^{\gamma j}|$ (d) for the photon+one-jet sample.*

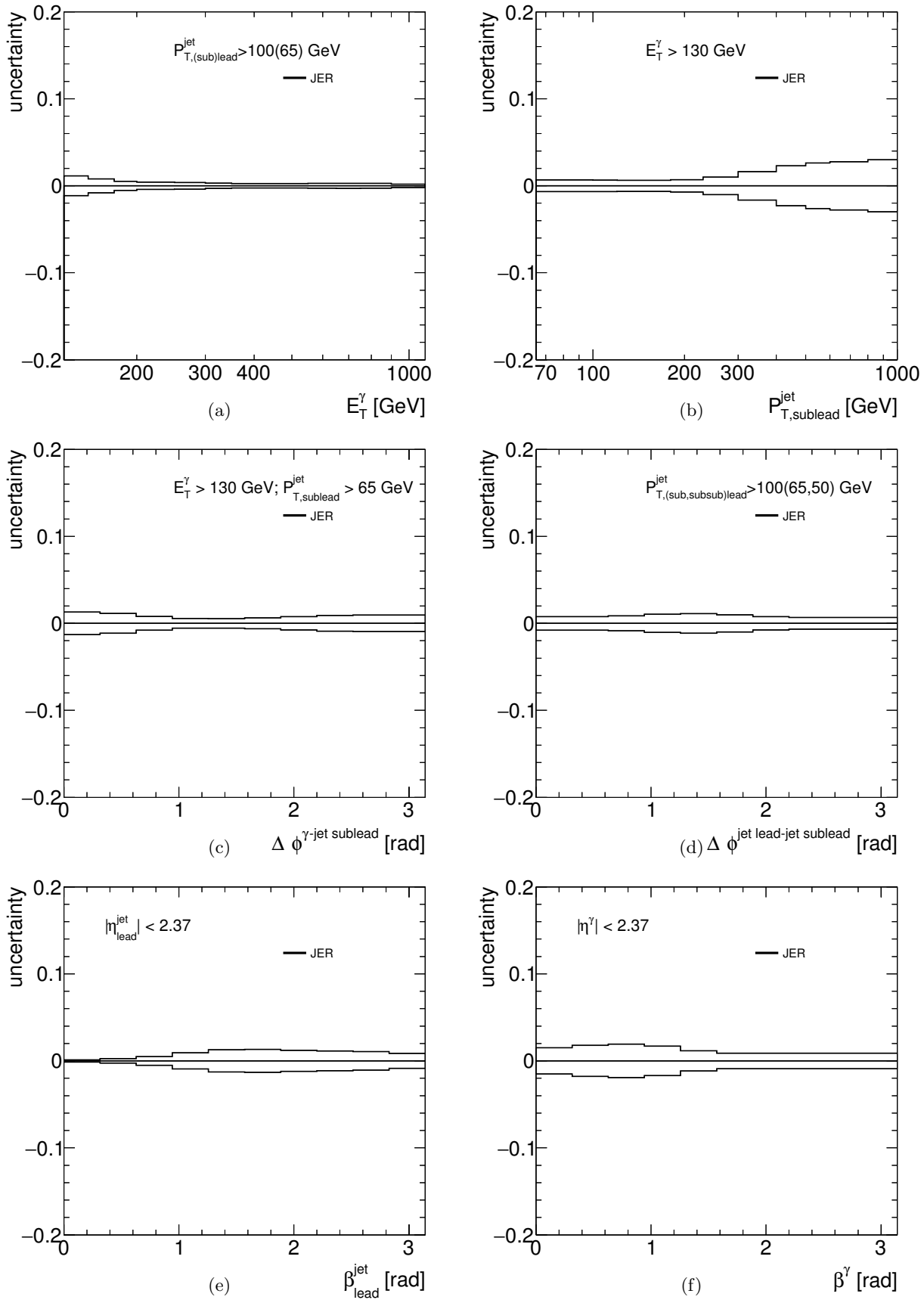


Figure 6.68: *Systematic uncertainties on the measured cross section due to the uncertainty in the jet energy resolution as functions of* E^{γ}_{T} *(a),* $p^{\mathrm{jet}}_{\mathrm{T,sublead}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$ *(d),* $\beta^{\mathrm{jet}}_{\mathrm{lead}}$ *(e) and* β^{γ} *(f) for the photon+two-jets (a,b,c,d) and the colour coherence (e,f) samples.*

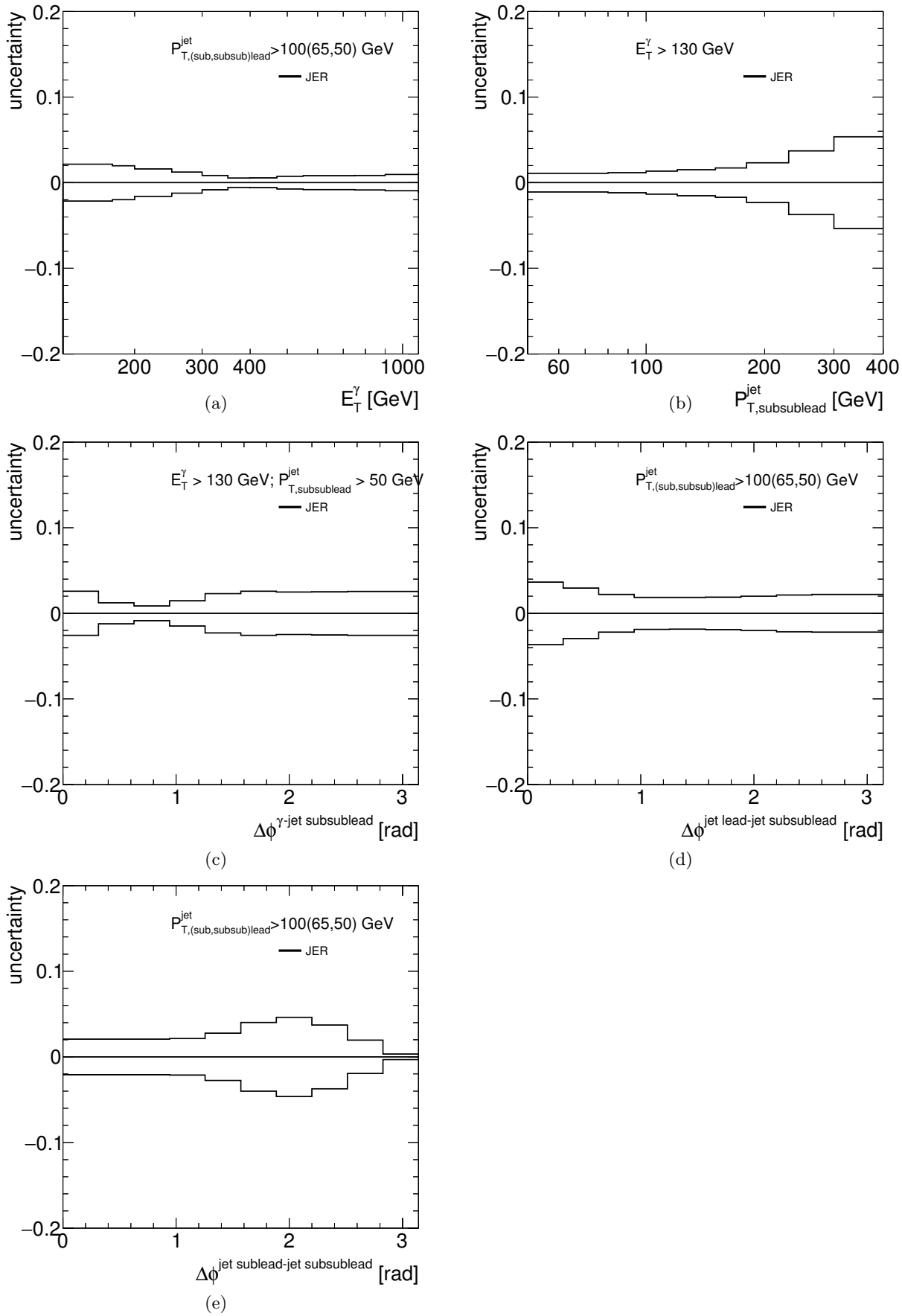


Figure 6.69: *Systematic uncertainties on the measured cross section due to the uncertainty in the jet energy resolution as functions of* E^{γ}_{T} *(a),* $p^{\mathrm{jet}}_{\mathrm{T,subsublead}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\ subsublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\ lead}-\mathrm{jet\ subsublead}}$ *(d),* $\Delta\phi^{\mathrm{jet\ sublead}-\mathrm{jet\ subsublead}}$ *(e) for the photon+three-jets sample.*

6.11.5 QCD cascade and hadronisation model dependence

The effect due to the QCD cascade and hadronisation models in the signal purity and correction factors was estimated as the deviations observed from the results of PYTHIA (nominal MC used for corrections) by using SHERPA MC. In the case of β^{γ} and $\beta^{\text{jet}}_{\text{lead}}$ SHERPA was used as the nominal MC and thus PYTHIA was used to estimate this uncertainty. Figures 6.70 to 6.72 show the resulting uncertainties as functions of the observables studied.

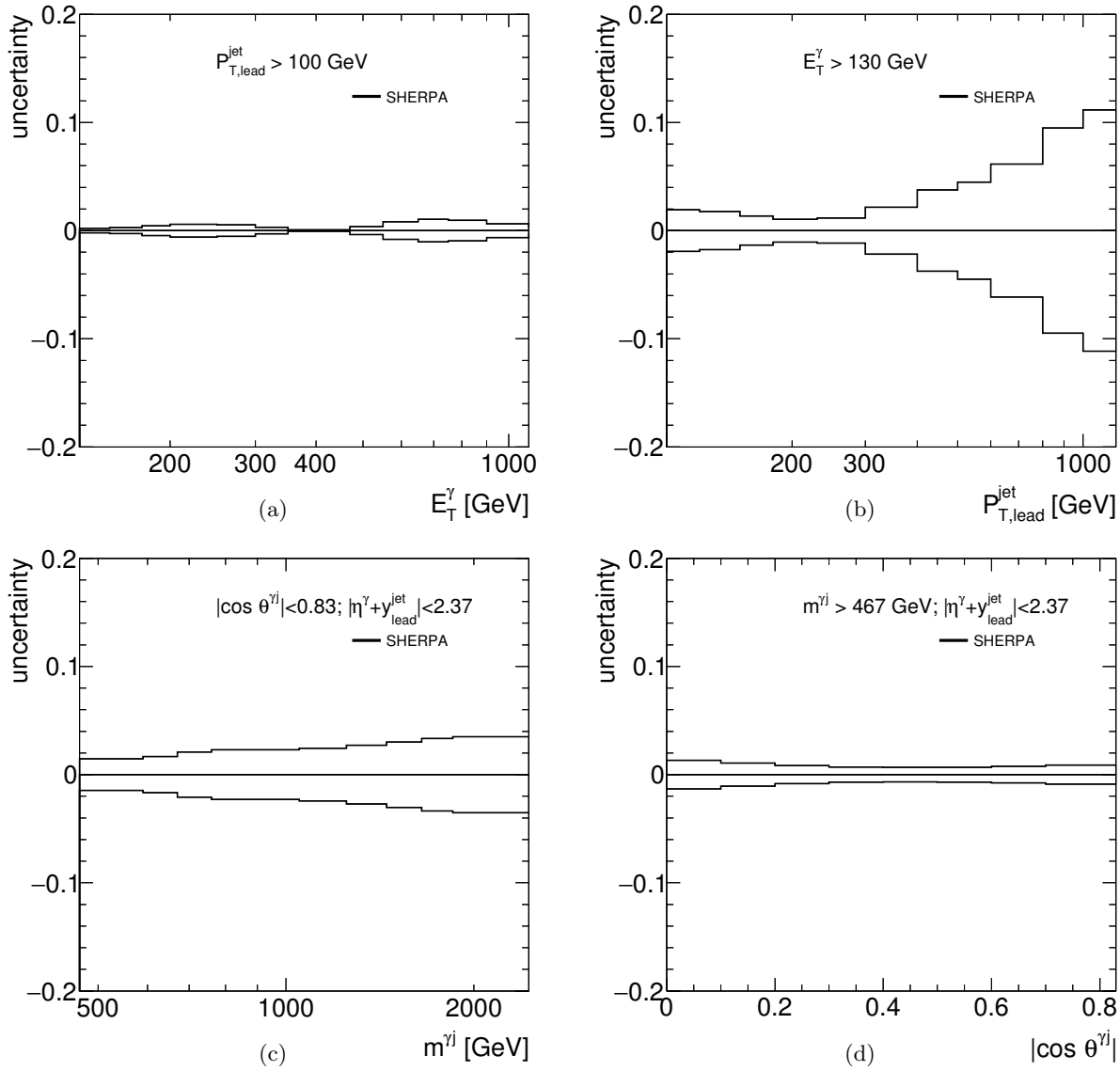


Figure 6.70: *Systematic uncertainties on the measured cross section due to the QCD cascade and the hadronisation model as functions of E^{γ}_{T} (a), $p^{\text{jet}}_{\text{T,lead}}$ (b), $m^{\gamma j}$ (c) and $|\cos\theta^{\gamma j}|$ (d) for the photon+one-jet sample.*

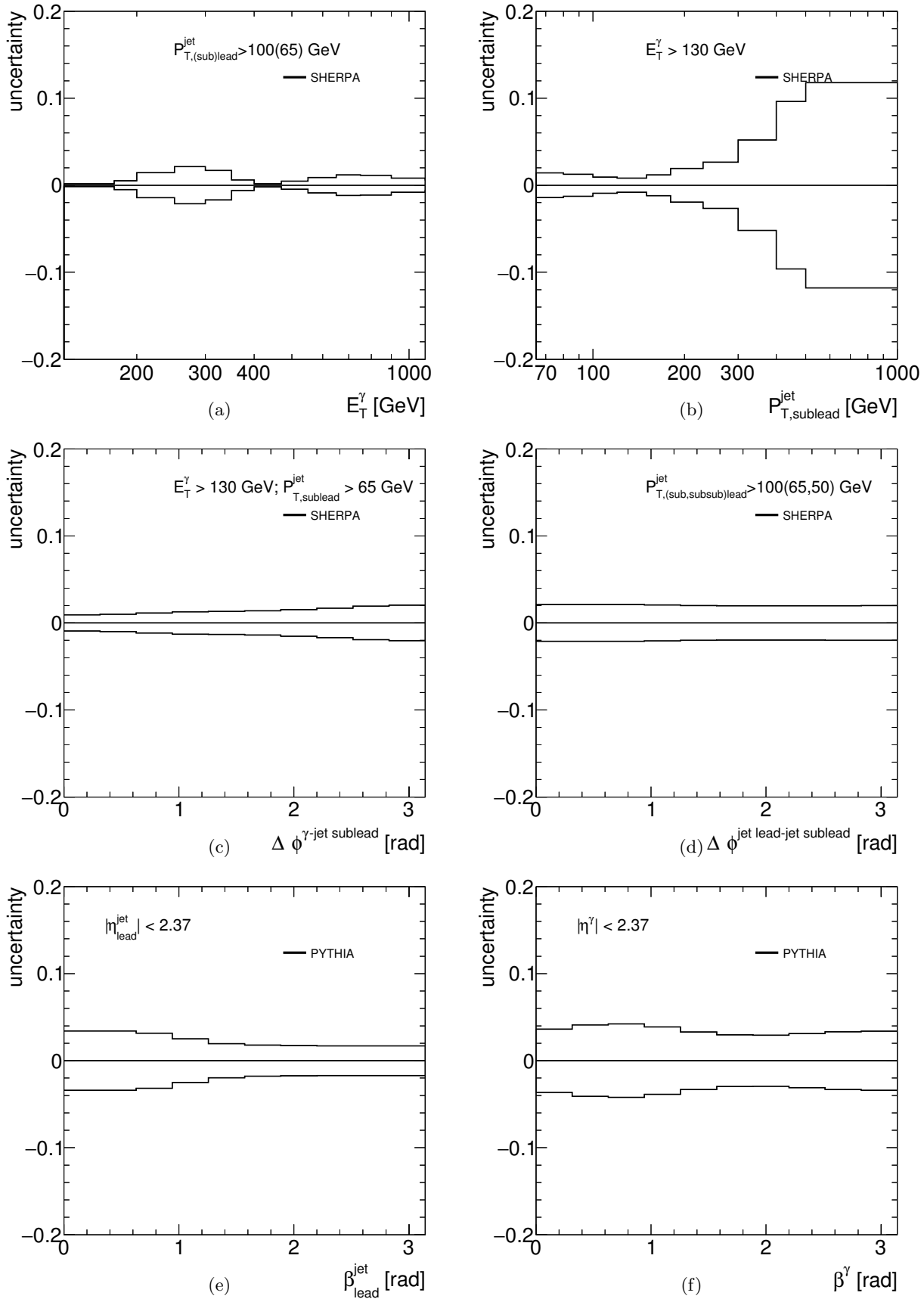


Figure 6.71: *Systematic uncertainties on the measured cross section due to the QCD cascade and the hadronisation model as functions of* E_{T}^{γ} *(a),* $p_{\mathrm{T,sublead}}^{\mathrm{jet}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$ *(d),* $\beta_{\mathrm{lead}}^{\mathrm{jet}}$ *(e) and* β^{γ} *(f) for the photon+two-jets (a,b,c,d) and the colour coherence (e,f) samples.*

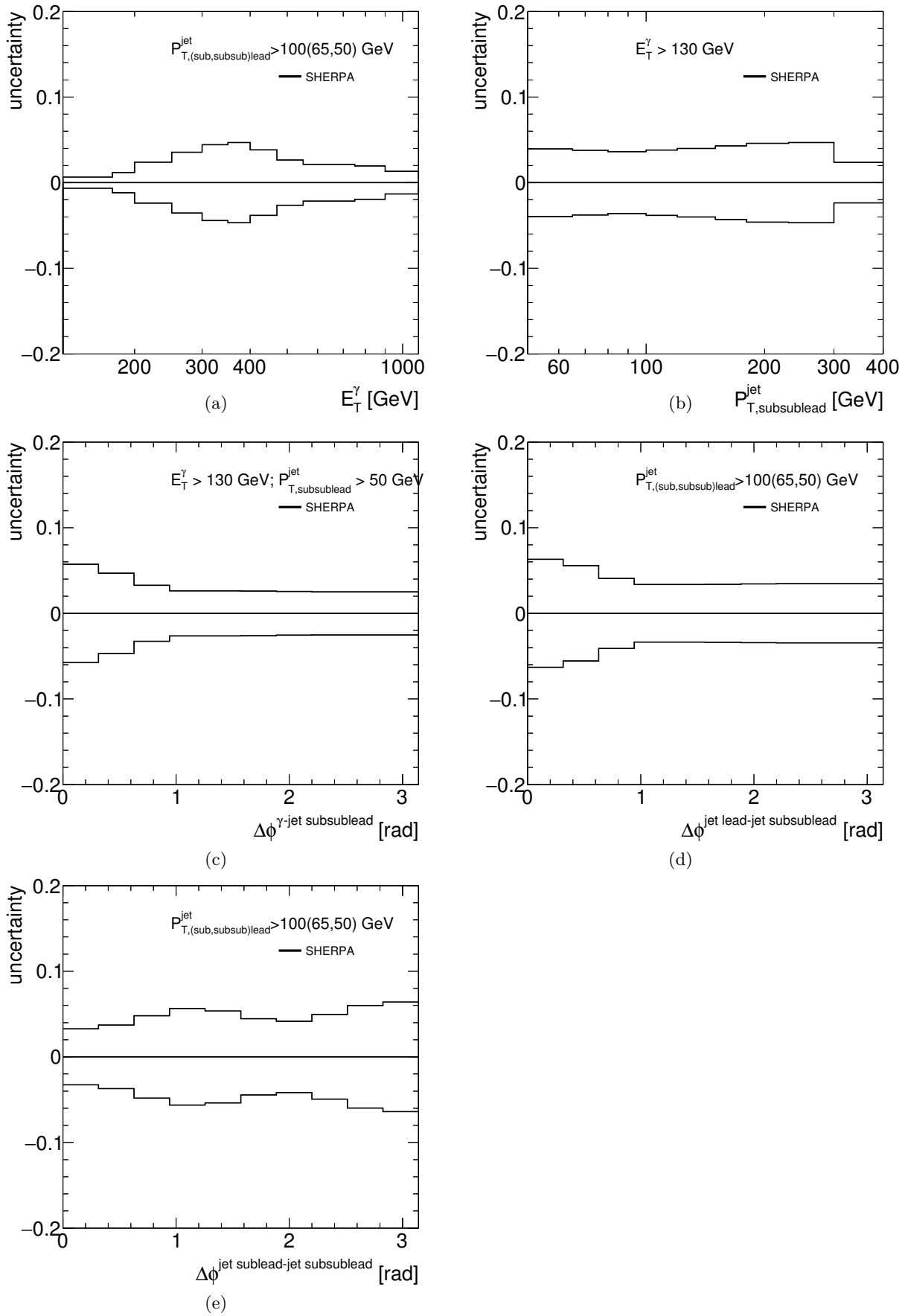


Figure 6.72: *Systematic uncertainties on the measured cross section due to the QCD cascade and the hadronisation model as functions of* E_{T}^{γ} *(a),* $p_{\mathrm{T,subsublead}}^{\mathrm{jet}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\ subsublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\ lead}-\mathrm{jet\ subsublead}}$ *(d),* $\Delta\phi^{\mathrm{jet\ sublead}-\mathrm{jet\ subsublead}}$ *(e) for the photon+three-jets sample.*

6.11.6 Photon identification efficiency

Scale factors were applied to the MC events to match the *tight* identification efficiency between data and simulation (see section 6.5). The uncertainty on the photon identification was estimated by propagating the uncertainty on these scale factors to the cross section. In addition, these scale factors were derived for a selection of $E_{\mathrm{T}}^{\mathrm{iso}} < 4$ GeV; therefore, an additional systematic uncertainty was included to account for the different $E_{\mathrm{T}}^{\mathrm{iso}}$ requirement in this analysis. This uncertainty was estimated from MC by taking the difference on the ID efficiency between the E_{T}^{γ}-dependent isolation and the $E_{\mathrm{T}}^{\mathrm{iso}} < 4$ GeV requirements. Figures 6.73 to 6.75 show the resulting uncertainties as functions of the observables studied.

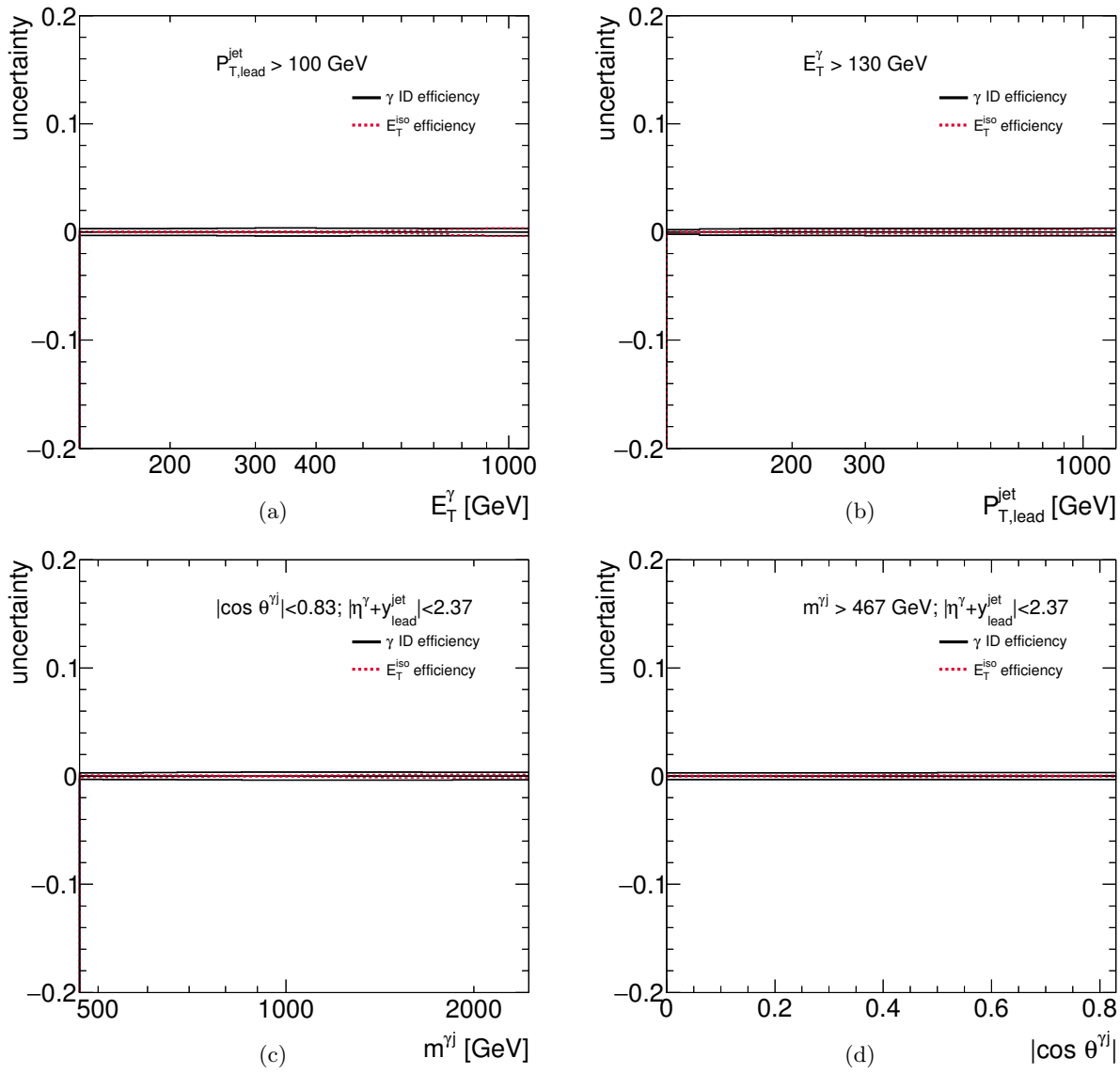


Figure 6.73: *Systematic uncertainties on the measured cross section due to the uncertaintiy in the photon ID efficiency as functions of* E_{T}^{γ} *(a),* $p_{\mathrm{T,lead}}^{\mathrm{jet}}$ *(b),* $m^{\gamma j}$ *(c) and* $|\cos\theta^{\gamma j}|$ *(d) for the photon+one-jet sample.*

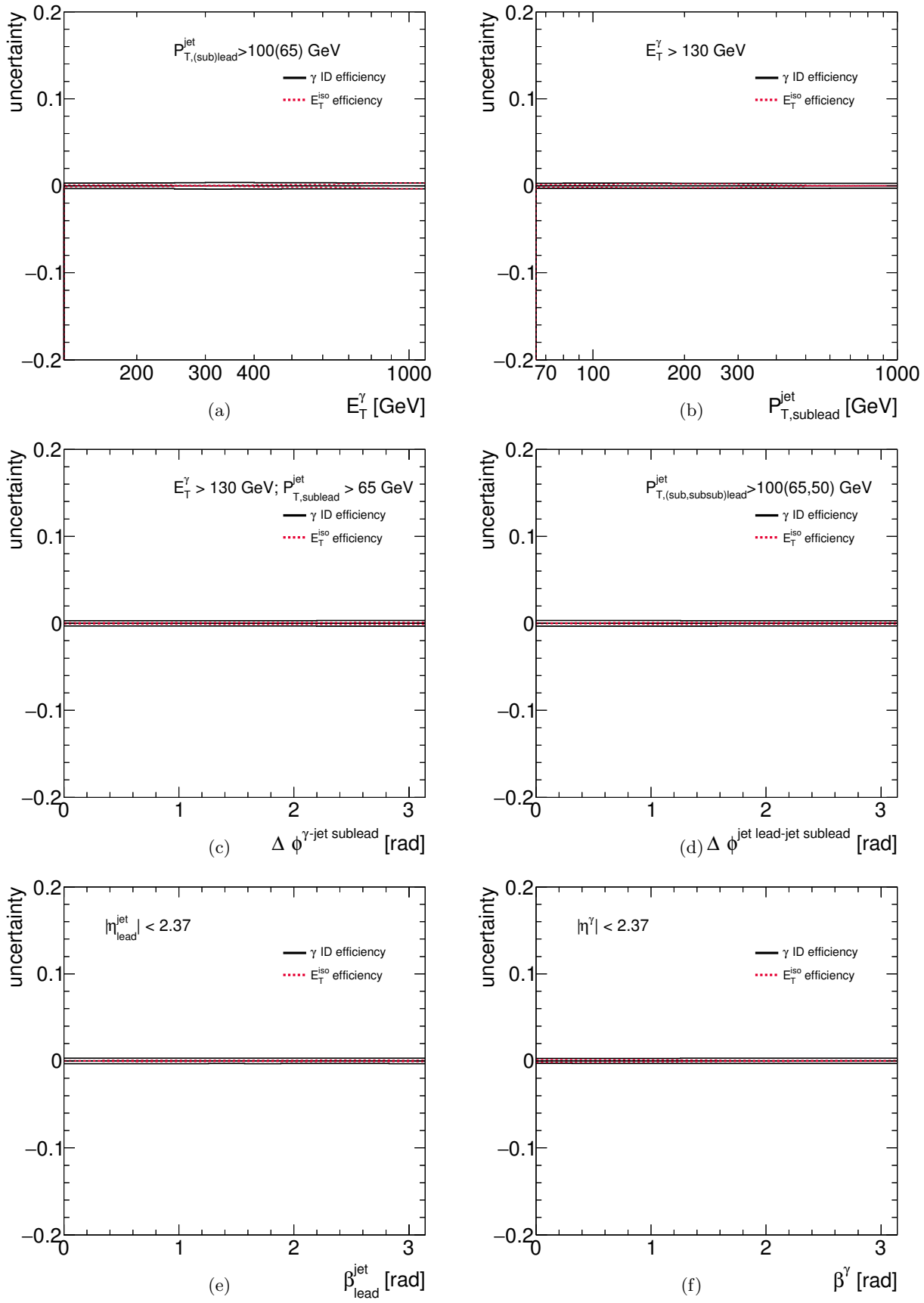


Figure 6.74: *Systematic uncertainties on the measured cross section due to the uncertaintiy in the photon ID efficiency as functions of* E^{γ}_{T} *(a),* $p^{jet}_{T,sublead}$ *(b),* $\Delta\phi^{\gamma-jet\ sublead}$ *(c),* $\Delta\phi^{jet\ lead-jet\ sublead}$ *(d),* β^{jet}_{lead} *(e) and* β^{γ} *(f) for the photon+two-jets (a,b,c,d) and the colour coherence (e,f) samples.*

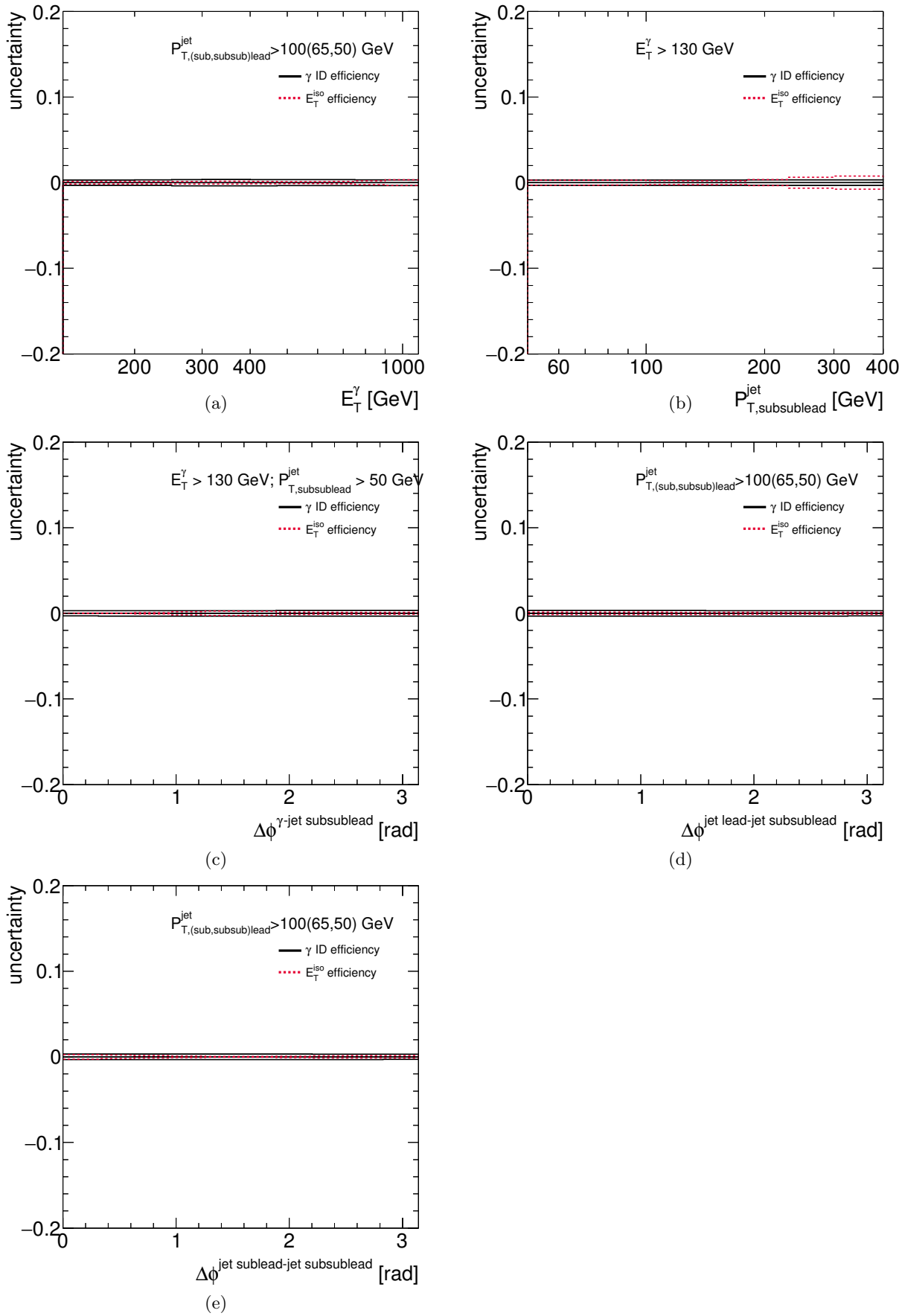


Figure 6.75: *Systematic uncertainties on the measured cross section due to the uncertaintiy in the photon ID efficiency as functions of* E_{T}^{γ} *(a),* $p_{\mathrm{T,subsublead}}^{\mathrm{jet}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\ subsublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\ lead}-\mathrm{jet\ subsublead}}$ *(d),* $\Delta\phi^{\mathrm{jet\ sublead}-\mathrm{jet\ subsublead}}$ *(e) for the photon+three-jets sample.*

6.11.7 Choice of background control regions

The 2D-sideband method (see section 6.6) was used to subtract the background in the signal region. The estimation of the background contamination in the signal region is affected by the choice of the background control regions.

The uncertainty due to this choice was estimated by varying the gap between regions A,C and B,D in the $E_{\mathrm{T}}^{\mathrm{iso}}$ variable (see figure 6.21), i.e., varying the E_{T}^{γ}-dependent isolation requirement for regions B and D ($E_{\mathrm{T}}^{\mathrm{iso}} > 4.8 + 4.2 \cdot 10^{-3} \times E_{\mathrm{T}}^{\gamma} + 2$ GeV) by 1 GeV up and down.

In addition, the choice of the inverted photon identification variables were varied. The nominal non-tight photon control region is defined by photons which pass *loose prime* but fail the *tight* identification criteria. The uncertainty due to this choice was estimated by repeating the analysis with a *tighter loose prime* and a *looser loose prime* (defined in section 4.4.2). Figures 6.76 to 6.78 show the resulting uncertainties as functions of the observables studied.

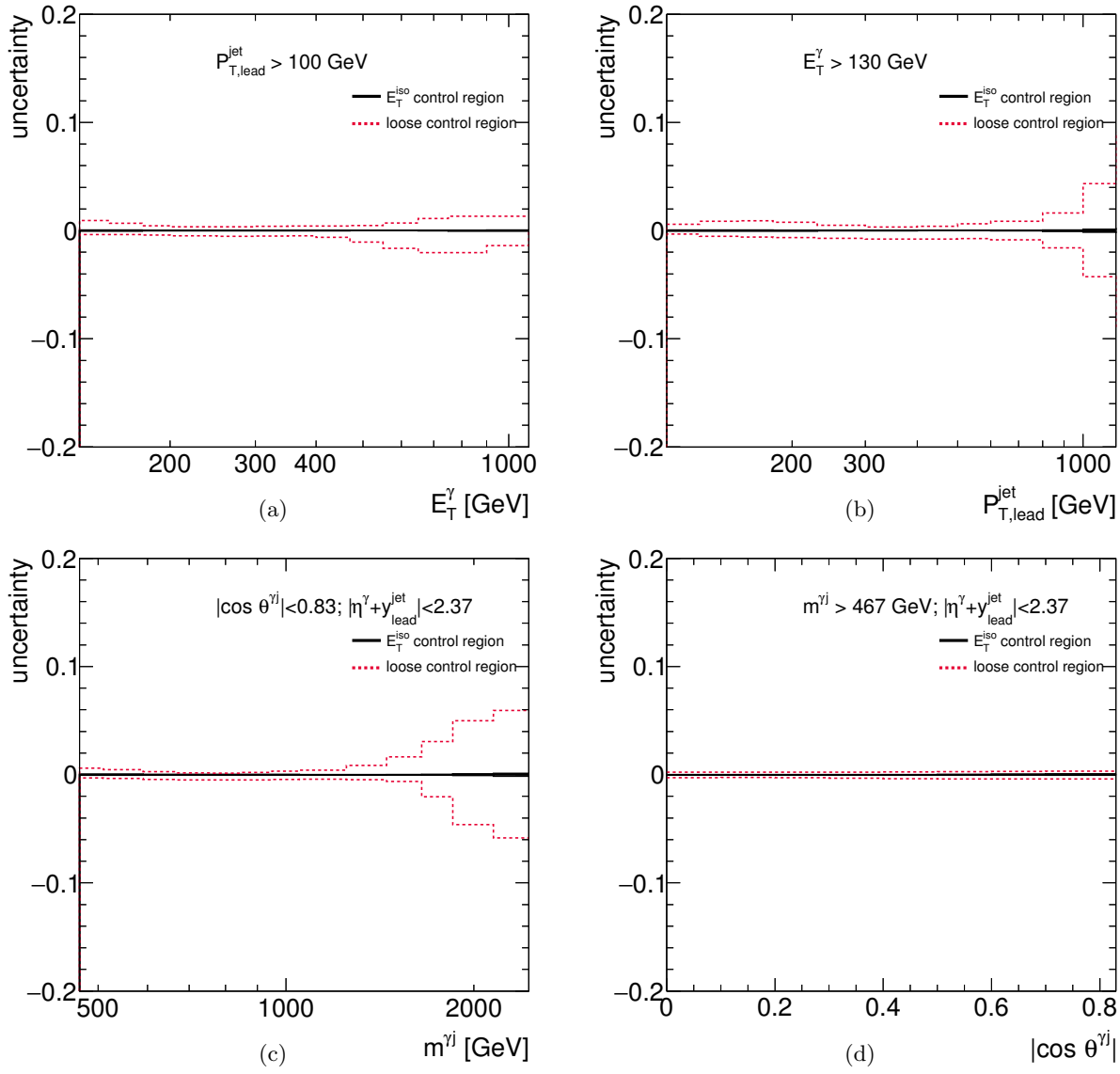


Figure 6.76: *Systematic uncertainties on the measured cross section due to the choice of background control regions as functions of* E_{T}^{γ} *(a),* $p_{\mathrm{T,lead}}^{\mathrm{jet}}$ *(b),* $m^{\gamma j}$ *(c) and* $|\cos\theta^{\gamma j}|$ *(d) for the photon+one-jet sample.*

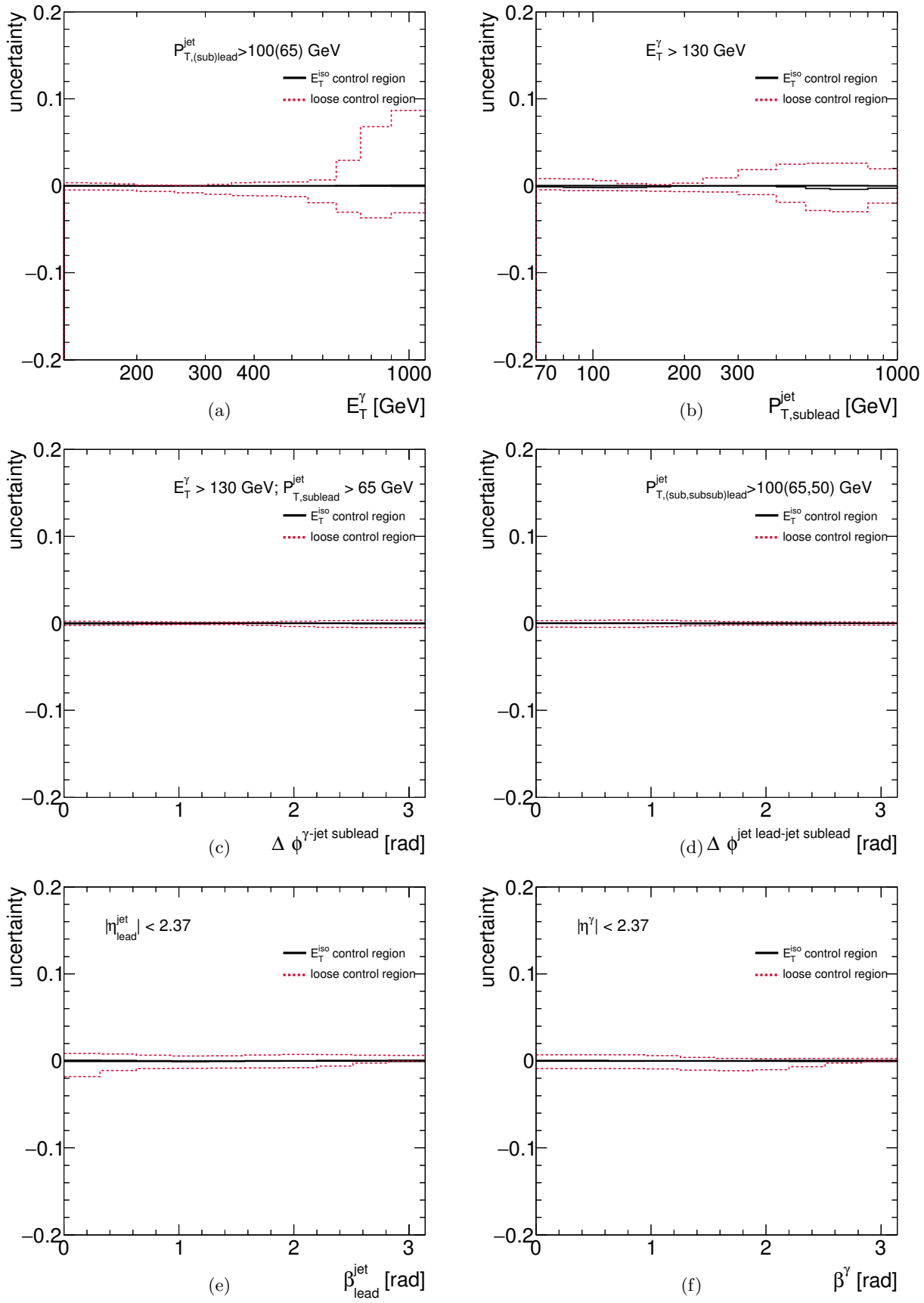


Figure 6.77: *Systematic uncertainties on the measured cross section due to the choice of background control regions as functions of* E_{T}^{γ} *(a),* $p_{\mathrm{T,sublead}}^{\mathrm{jet}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$ *(d),* $\beta_{\mathrm{lead}}^{\mathrm{jet}}$ *(e) and* β^{γ} *(f) for the photon+two-jets (a,b,c,d) and the colour coherence (e,f) samples.*

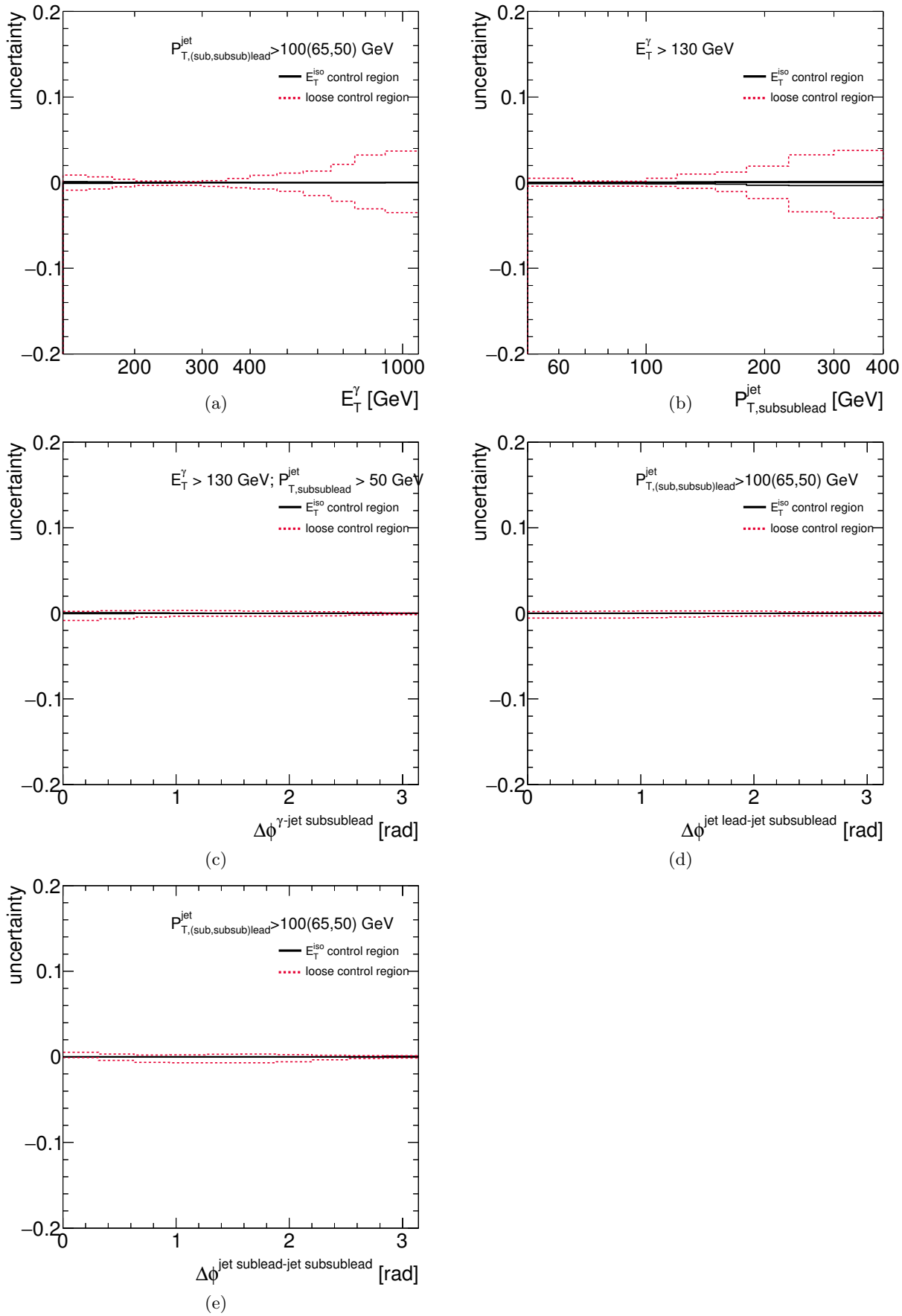


Figure 6.78: *Systematic uncertainties on the measured cross section due to the choice of background control regions as functions of* E^{γ}_{T} *(a),* $p^{\mathrm{jet}}_{\mathrm{T,subsublead}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\ subsublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\ lead}-\mathrm{jet\ subsublead}}$ *(d),* $\Delta\phi^{\mathrm{jet\ sublead}-\mathrm{jet\ subsublead}}$ *(e) for the photon+three-jets sample.*

6.11.8 Signal modelling

The MC simulation of the signal was used to estimate the signal leakage fraction in the 2D side band method (see section 6.6) and to compute the bin-by-bin correction factors.

For the signal leakage fractions, the PYTHIA simulation was used with the admixture of the hard and brem components as given by the MC to yield the background substracted data distributions. The signal leakage fractions depend on the relative fraction of the two components. The uncertainty related to the simulation of the PYTHIA components in the signal leakage fractions was estimated by performing the background subtraction with the optimised admixture for each observable (see section 6.8).

To compute the bin-by-bin correction factors, the relative fraction of the components in PYTHIA were first fitted to the background-subtracted data distributions (see section 6.8). A systematic uncertainty due to the fit was estimated using the error from the fit, i.e. α was replaced by $\alpha \pm \Delta\alpha$ in equation 6.5 to add the components.

These two uncertainties were not computed for β^{γ} and $\beta^{\text{jet}}_{\text{lead}}$ since SHERPA was used as the nominal MC both to compute the bin-by-bin correction factors and the leakage fractions.

Figures 6.79 to 6.81 show the resulting uncertainties as functions of the observables estudied.

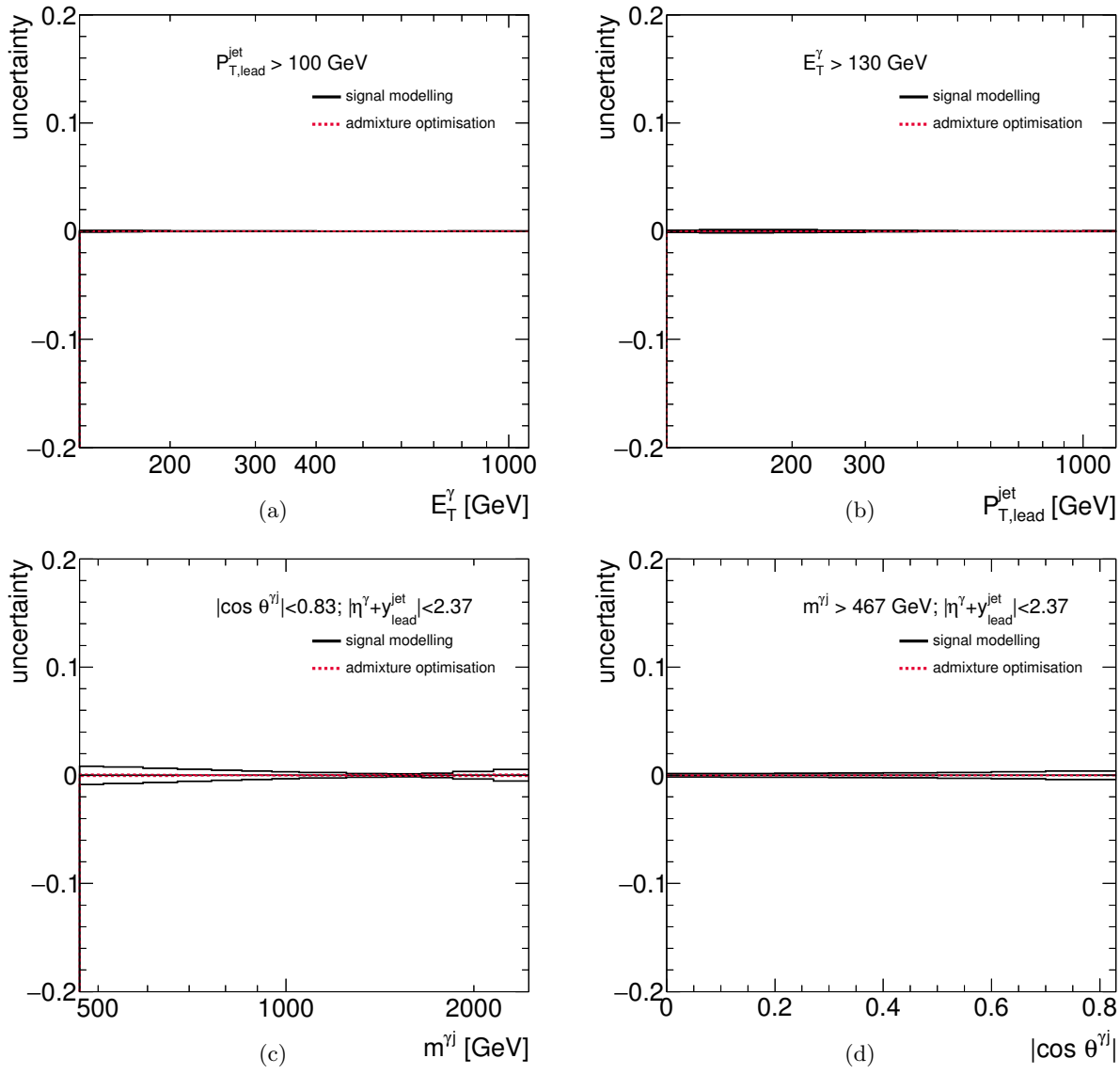


Figure 6.79: *Systematic uncertainties on the measured cross section due to the effect of the signal modelling in the signal purity and in the bin-by-bin correction factors as functions of E_{T}^{γ} (a), $p_{\mathrm{T,lead}}^{\mathrm{jet}}$ (b), $m^{\gamma j}$ (c) and $|\cos\theta^{\gamma j}|$ (d) for the photon+one-jet sample.*

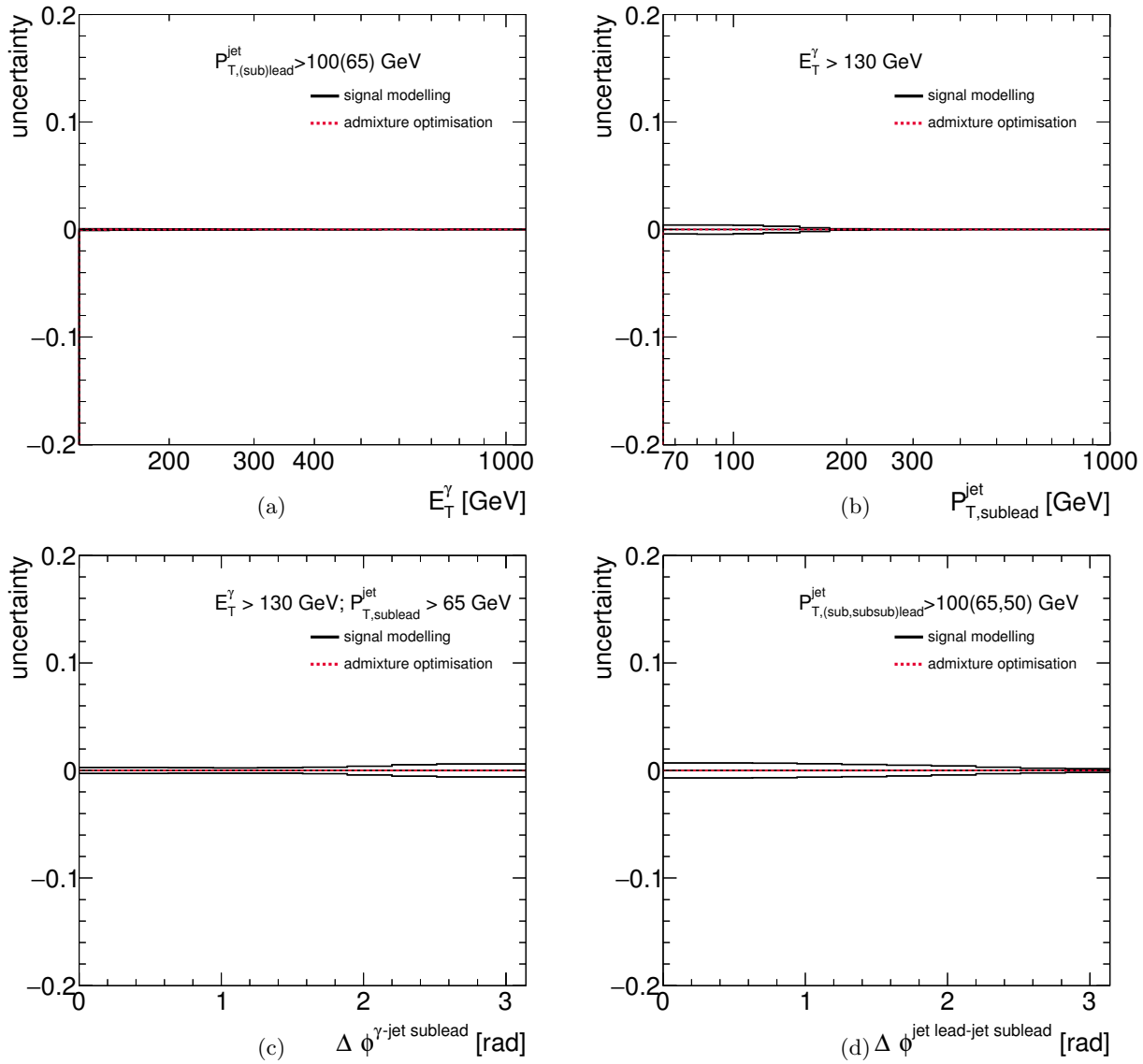


Figure 6.80: *Systematic uncertainties on the measured cross section due to the effect of the signal modelling in the signal purity and in the bin-by-bin correction factors as functions of* E^{γ}_{T} *(a),* $p^{\mathrm{jet}}_{\mathrm{T,sublead}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\ sublead}}$ *(c) and* $\Delta\phi^{\mathrm{jet\ lead}-\mathrm{jet\ sublead}}$ *(d) for the photon+two-jets sample.*

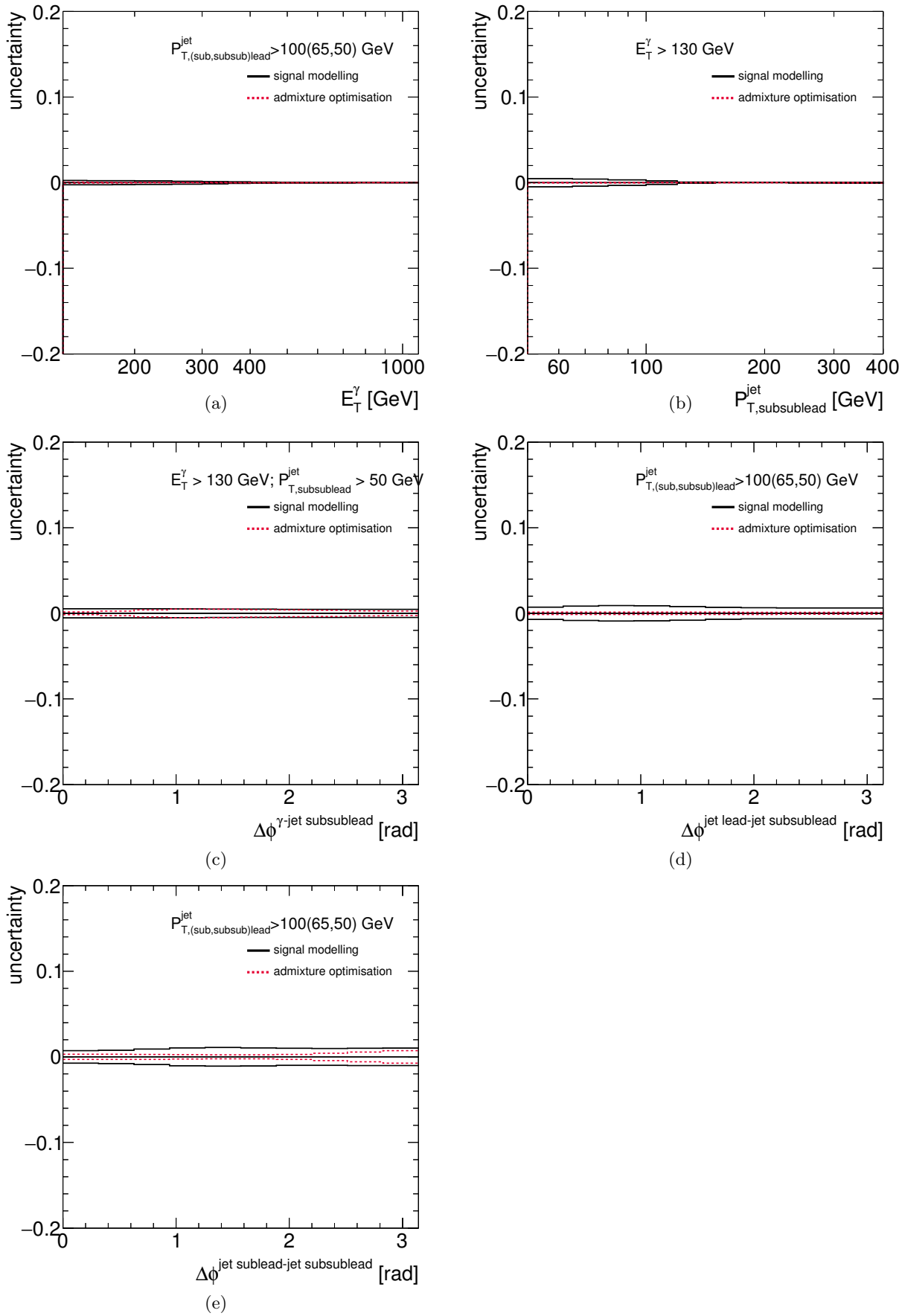


Figure 6.81: *Systematic uncertainties on the measured cross section due to the effect of the signal modelling in the signal purity and in the bin-by-bin correction factors as functions of* E_{T}^{γ} *(a),* $p_{\mathrm{T,subsublead}}^{\mathrm{jet}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\ subsublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\ lead}-\mathrm{jet\ subsublead}}$ *(d),* $\Delta\phi^{\mathrm{jet\ sublead}-\mathrm{jet\ subsublead}}$ *(e) for the photon+three-jets sample.*

6.11.9 Identification and isolation correlation in the background

The isolation and identification used to define the plane in the 2D side-band method to subtract the background (see section 6.6) were assumed to be uncorelated for background events ($R_{\mathrm{bg}} = 1$ in equation 6.3). Any correlation between these variables would affect the estimation of the purity of the signal and would lead to systematic uncertainties in the background subtraction procedure. A $\pm 10\%$ uncertainty was taken, varying R_{bg} to 0.9 and 1.1 respectively. Figures 6.82 to 6.84 show the resulting uncertainties as functions of the observables estudied.

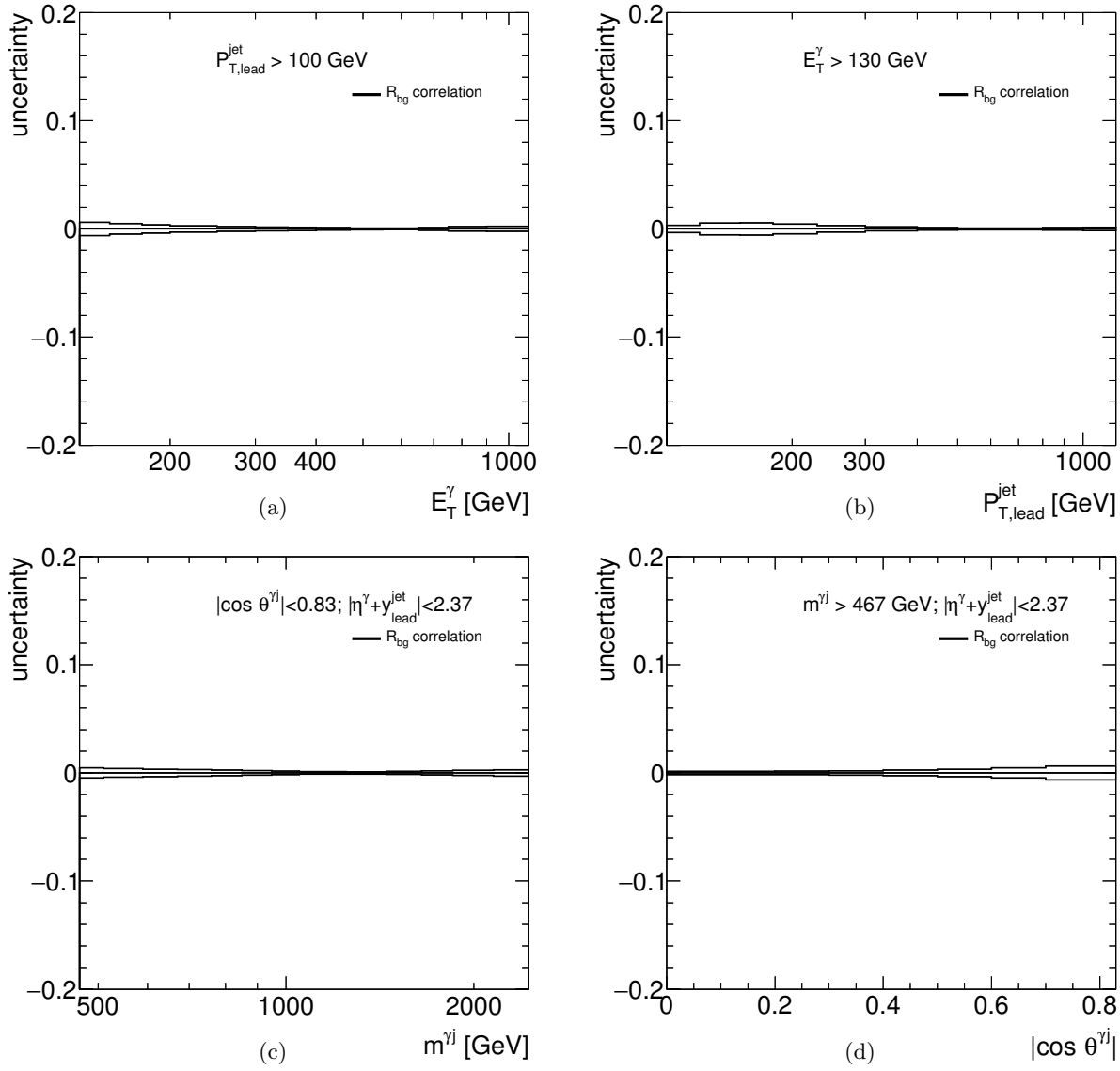


Figure 6.82: *Systematic uncertainties on the measured cross section due to the background identification and isolation correlation as functions of E^{γ}_{T} (a), $p^{\mathrm{jet}}_{\mathrm{T,lead}}$ (b), $m^{\gamma j}$ (c) and $|\cos\theta^{\gamma j}|$ (d) for the photon+one-jet sample.*

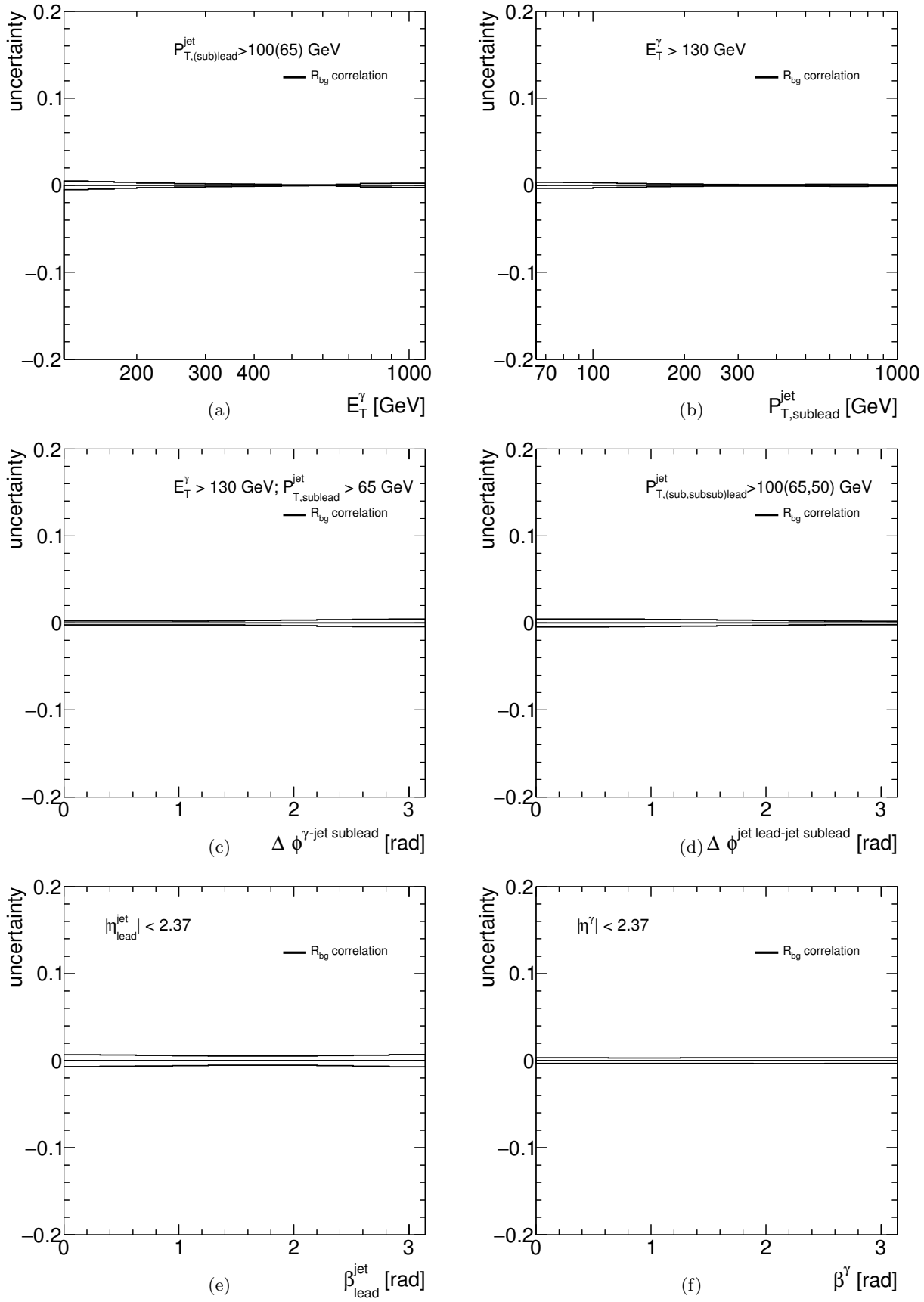


Figure 6.83: *Systematic uncertainties on the measured cross section due to the background identification and isolation correlation as functions of* E_{T}^{γ} *(a),* $p_{\mathrm{T,sublead}}^{\mathrm{jet}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$ *(d),* $\beta_{\mathrm{lead}}^{\mathrm{jet}}$ *(e) and* β^{γ} *(f) for the photon+two-jets (a,b,c,d) and the colour coherence (e,f) samples.*

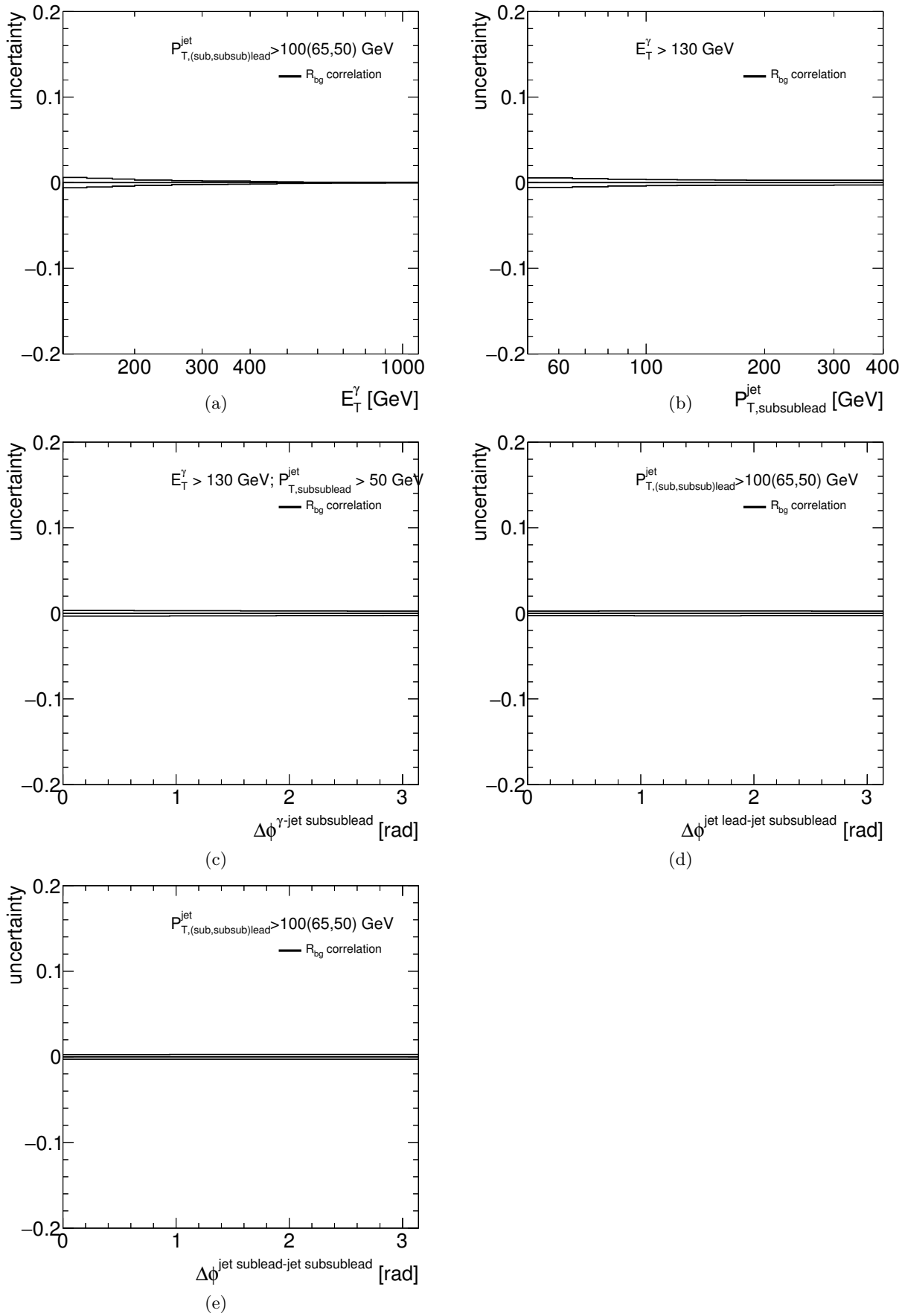


Figure 6.84: *Systematic uncertainties on the measured cross section due to the background identification and isolation correlation as functions of* E_{T}^{γ} *(a),* $p_{\mathrm{T,subsublead}}^{\mathrm{jet}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\ subsublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\ lead}-\mathrm{jet\ subsublead}}$ *(d),* $\Delta\phi^{\mathrm{jet\ sublead}-\mathrm{jet\ subsublead}}$ *(e) for the photon+three-jets sample.*

6.11.10 MC sample statistics

The limited MC statistics mainly affects the bin-by-bin corretion factors. Figures 6.85 to 6.87 show the statistical uncertainty of the MC samples as functions of the observables studied.

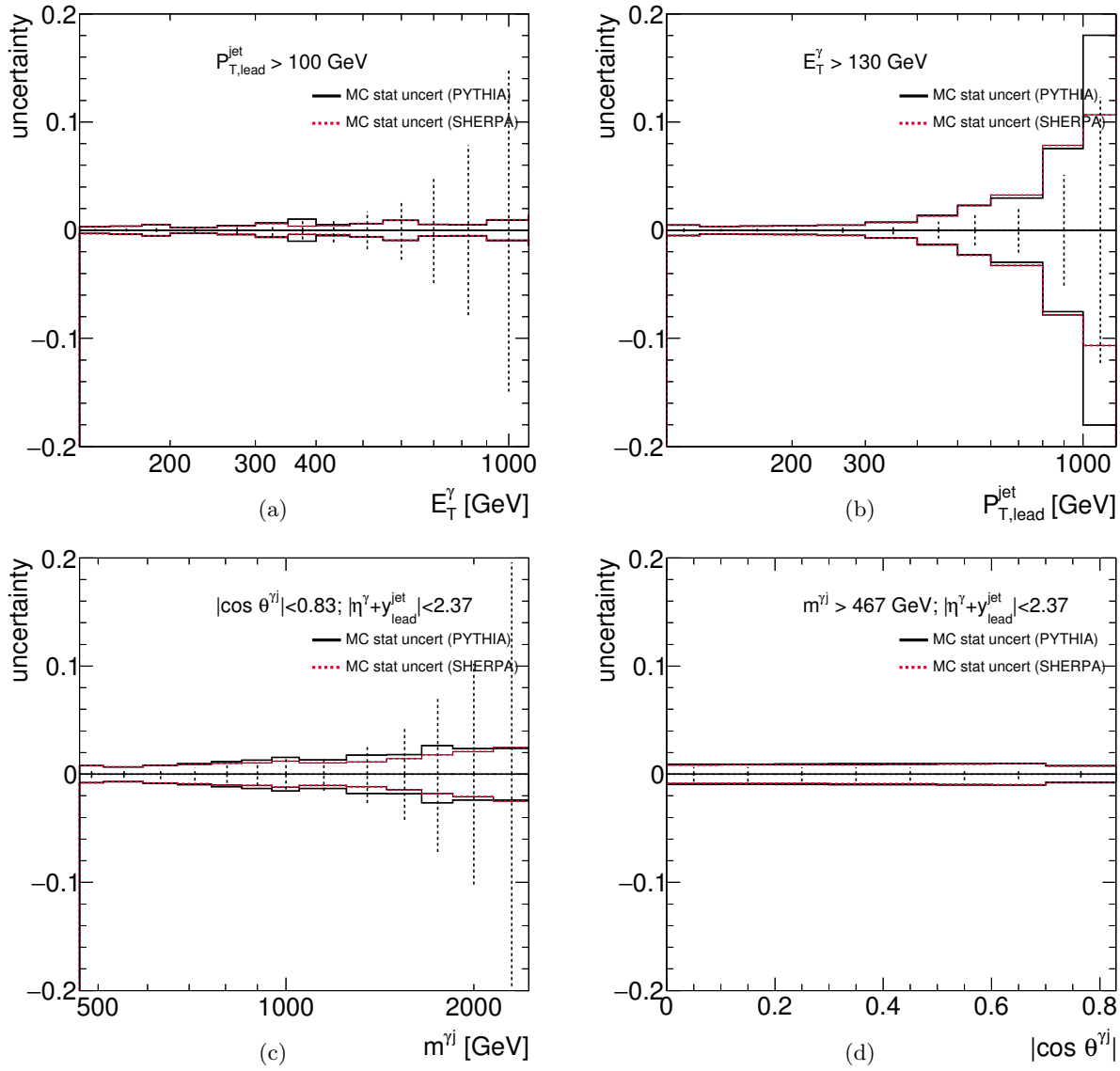


Figure 6.85: *Statistical uncertainty of the MC samples as a function of E^{γ}_{T} (a), $p^{\mathrm{jet}}_{\mathrm{T,lead}}$ (b), $m^{\gamma j}$ (c) and $|\cos\theta^{\gamma j}|$ (d) for the photon+one-jet sample. For comparison, the vertical error bars display the data statistical uncertainty.*

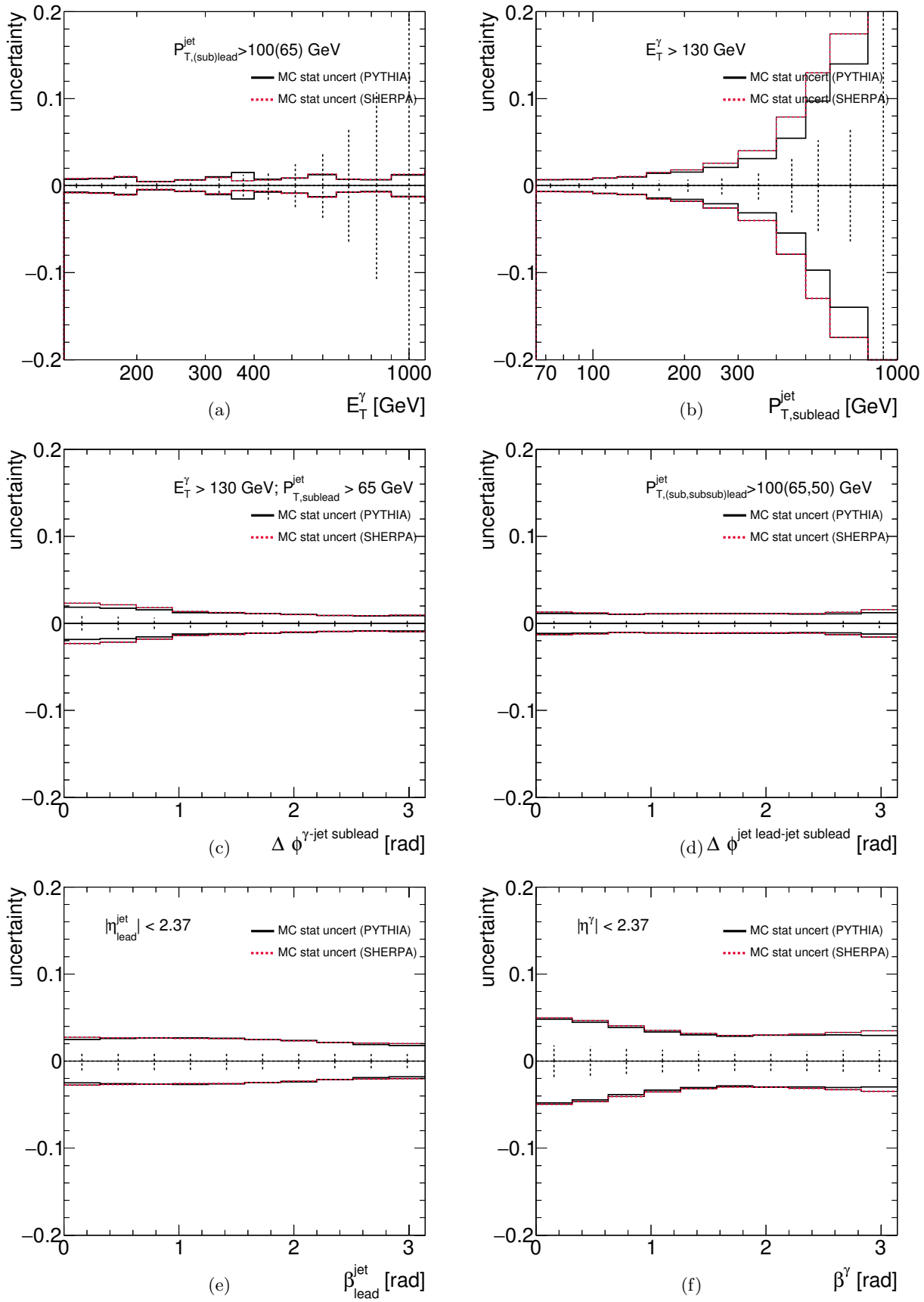


Figure 6.86: *Statistical uncertainty of the MC samples as a function of* E_{T}^{γ} *(a),* $p_{\mathrm{T,sublead}}^{\mathrm{jet}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$ *(d),* $\beta_{\mathrm{lead}}^{\mathrm{jet}}$ *(e) and* β^{γ} *(f) for the photon+two-jets (a,b,c,d) and the colour coherence (e,f) samples. For comparison, the vertical error bars display the data statistical uncertainty.*

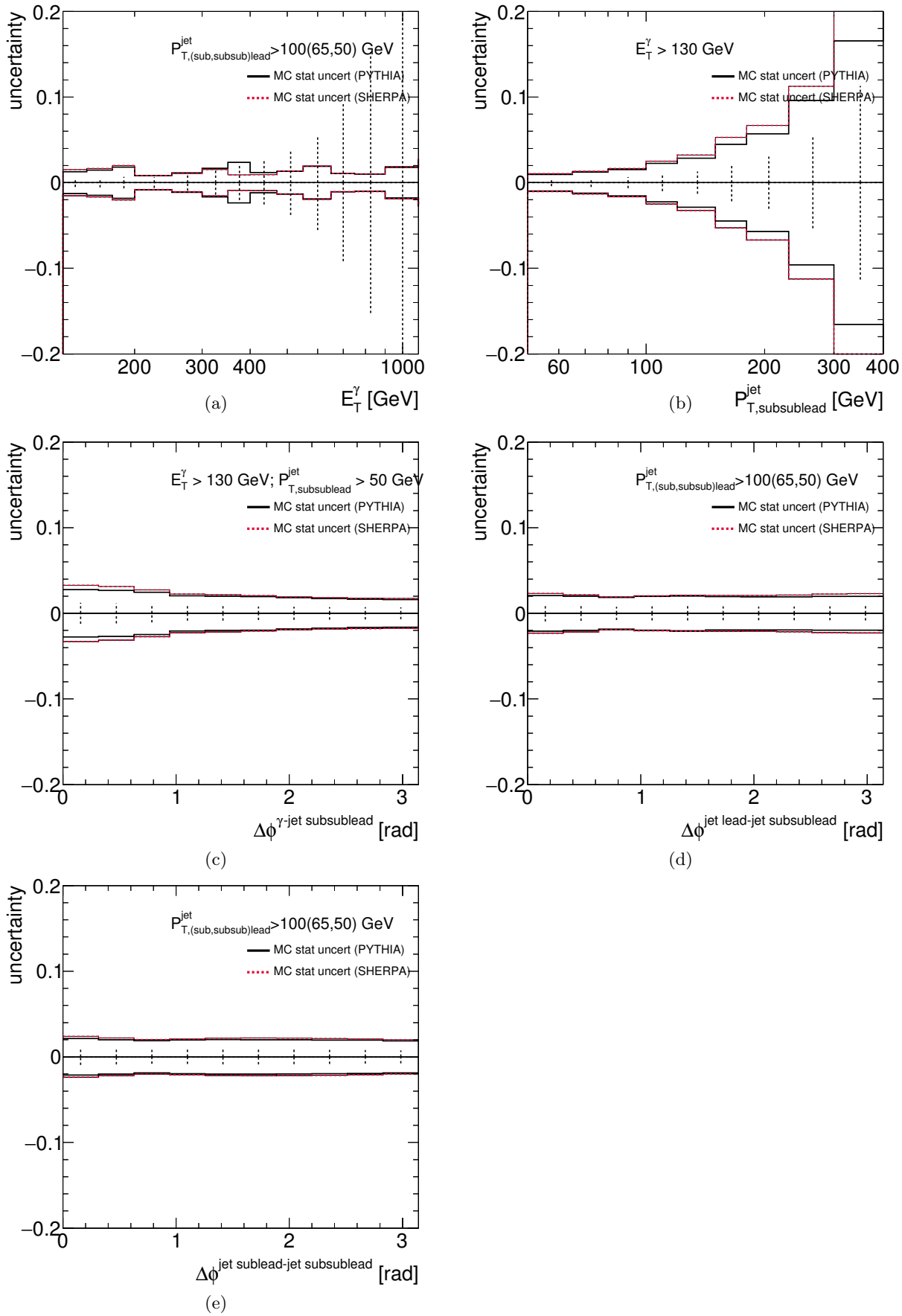


Figure 6.87: *Statistical uncertainty of the MC samples as a function of* E^{γ}_{T} *(a),* $p^{\mathrm{jet}}_{\mathrm{T,subsublead}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\ subsublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\ lead}-\mathrm{jet\ subsublead}}$ *(d),* $\Delta\phi^{\mathrm{jet\ sublead}-\mathrm{jet\ subsublead}}$ *(e) for the photon+three-jets sample. For comparison, the vertical error bars display the data statistical uncertainty.*

6.11.11 Total uncertainty

The total systematic uncertainty was computed by adding in quadrature the sources of uncertainty listed in the previous sections. Figures 6.88 to 6.90 show the resulting total uncertainties, the total systematic uncertainty added in quadrature with the statistical uncertainty, as a function of the observables studied. Statistical uncertainties for the data and MC are also shown separately.

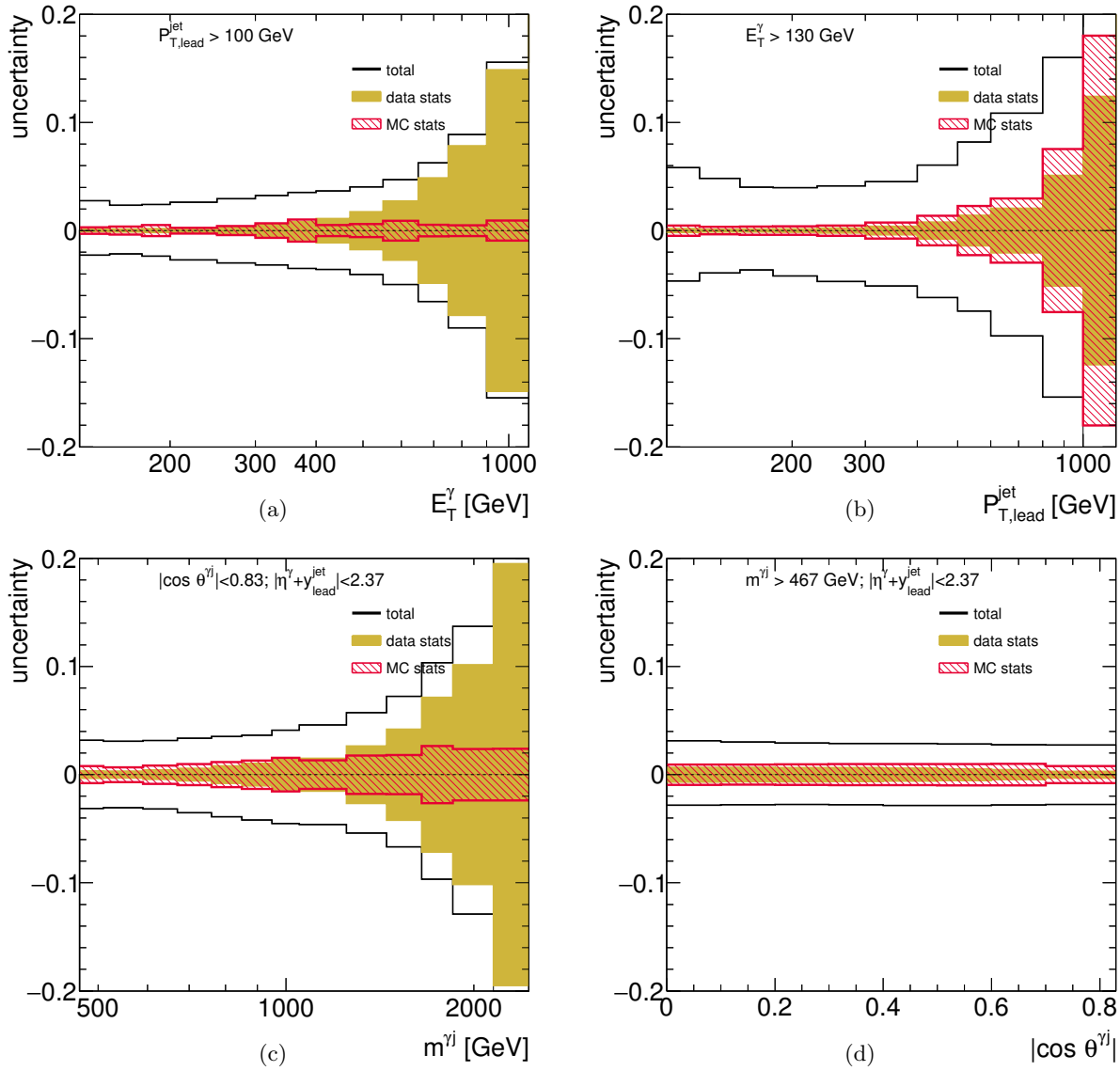


Figure 6.88: *Total uncertainties (white area) and statistical uncertainties from the data (shaded area) and MC (hatched area) as functions of* E^{γ}_{T} *(a),* $p^{\mathrm{jet}}_{\mathrm{T,lead}}$ *(b),* $m^{\gamma j}$ *(c) and* $|\cos\theta^{\gamma j}|$ *(d) for the photon+one-jet sample.*

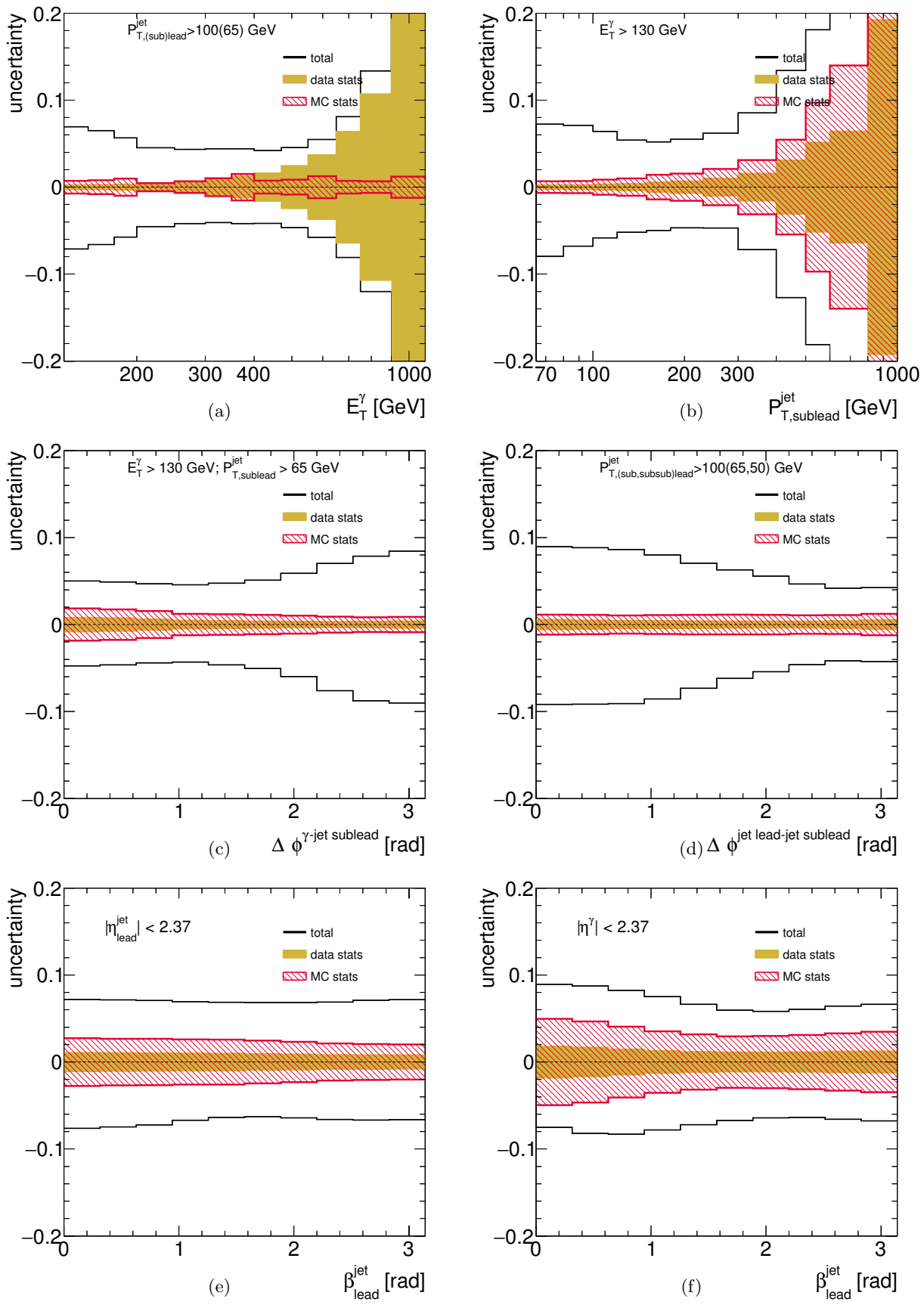


Figure 6.89: *Total uncertainties (white area) and statistical uncertainties from the data (shaded area) and MC (hatched area) as functions of* E^{γ}_{T} *(a),* $p^{\mathrm{jet}}_{\mathrm{T,sublead}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$ *(d),* $\beta^{\mathrm{jet}}_{\mathrm{lead}}$ *(e) and* β^{γ} *(f) for the photon+two-jets (a,b,c,d) and the colour coherence (e,f) samples.*

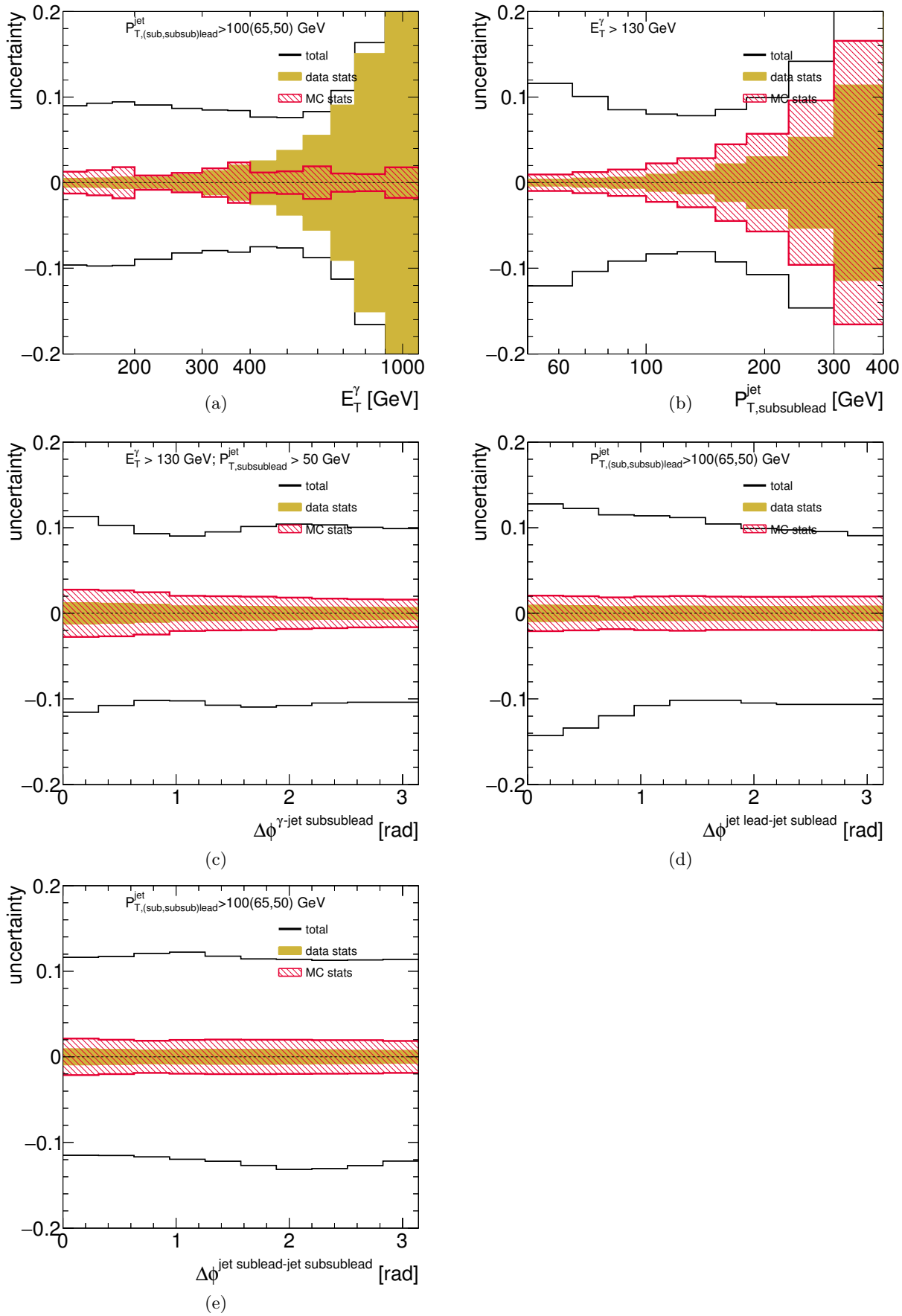


Figure 6.90: *Total uncertainties (white area) and statistical uncertainties from the data (shaded area) and MC (hatched area) as functions of* E_{T}^{γ} *(a),* $p_{\mathrm{T,subsublead}}^{\mathrm{jet}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\ subsublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\ lead}-\mathrm{jet\ subsublead}}$ *(d),* $\Delta\phi^{\mathrm{jet\ sublead}-\mathrm{jet\ subsublead}}$ *(e) for the photon+three-jets sample.*

6.11.12 Additional sources of systematic uncertainty

Other possible sources of systematic uncertainty were identified and investigated. Since their effect on the measured cross sections was determined to be negligible, they are not included in the total systematic uncertainty described above and are only listed in this section.

Unfolding method

A cross check of the bin-by-bin unfolding method was performed by using a bayesian unfolding (see section 7.4). The differences in the cross sections computed by the two methods are typically much smaller than 1% except in regions where the MC statistics is poor and so the bayesian unfolding is less trustworthy than the bin-by-bin estimation (see figures 7.32 to 7.34). This uncertainty was neglected.

Trigger uncertainty

The trigger efficiency in data was investigated (see section 7.2) and found to be 99.87% for tight and isolated photons with $E_{\mathrm{T}}^{\gamma} > 130$ GeV, whereas in MC it was taken as 100%. The difference in this efficiency between data and MC would consitute a systematic uncertainty, but it is so small that it was neglected.

Electron fake rate

A study of the rate of electrons faking a photon signal has been performed [91] using MC samples of $Z \to e^+e^-$ and $W \to e\nu$. It shows that the fake rate in the region of interest for this analysis ($E_{\mathrm{T}}^{\gamma} > 130$ GeV) is very small, so any correction, subtraction or uncertainty was neglected.

Pile-up

The pile-up interactions can have an effect on the measurements through differences between data and MC at high and low μ, differences in the signal purity and in the acceptance correction factors. Several tests were performed (see section 7.5) and no significant effect was observed.

6.12 Next-to-leading order QCD calculations for photon plus one jet

The NLO QCD calculations used in the photon+one-jet analysis presented here were computed using the program JETPHOX 1.3.2 [92, 93]. This program includes a full NLO QCD calculation of both the direct-photon and fragmentation contributions to the cross section for the $pp \to \gamma + 1\mathrm{Jet}$ reaction (see section 1.5 for the relevant theoretical framework).

The number of flavours was set to five. The renormalisation (μ_R), factorisation (μ_F) and fragmentation (μ_f) scales were chosen to be $\mu_R = \mu_F = \mu_f = E_{\mathrm{T}}^{\gamma}$. The calculations were done using the NLO CT10 parametrisations of the proton PDFs and the NLO photon BFG set II photon fragmentation function [94]. The strong coupling constant was calculated at two loops with $\alpha_s(m_Z) = 0.118$. Predictions based on the MSTW2008nlo [95] parametrisations of the proton PDFs were also computed.

The calculations were performed using a parton-level isolation cut, which required a total transverse energy below 10 GeV to mimic the particle level isolation cut (see section 6.3.1). $E_{\mathrm{T}}^{\mathrm{iso}}$

was calculated using the partons inside a cone of radious $R = 0.4$ around the photon direction. The anti-k_T algorithm with distance parameter $R = 0.6$ was applied to the partons in the events generated by JETPHOX to compute the cross-section predictions.

Figure 6.91 shows the predicted cross sections for the photon+one-jet observables. Shown are the calculations based on the CT10 and MSTW2008nlo parametrisations of the proton PDFs. The predictions based on MSTW2008nlo have similar shape to those based on CT10, but have $\sim 5\%$ larger normalisation.

The parton level predictions of PYTHIA are compared to the NLO QCD calculations based on CT10 in figure 6.92. The parton-level cross sections of PYTHIA give an adequate description of the shape of the NLO QCD predictions.

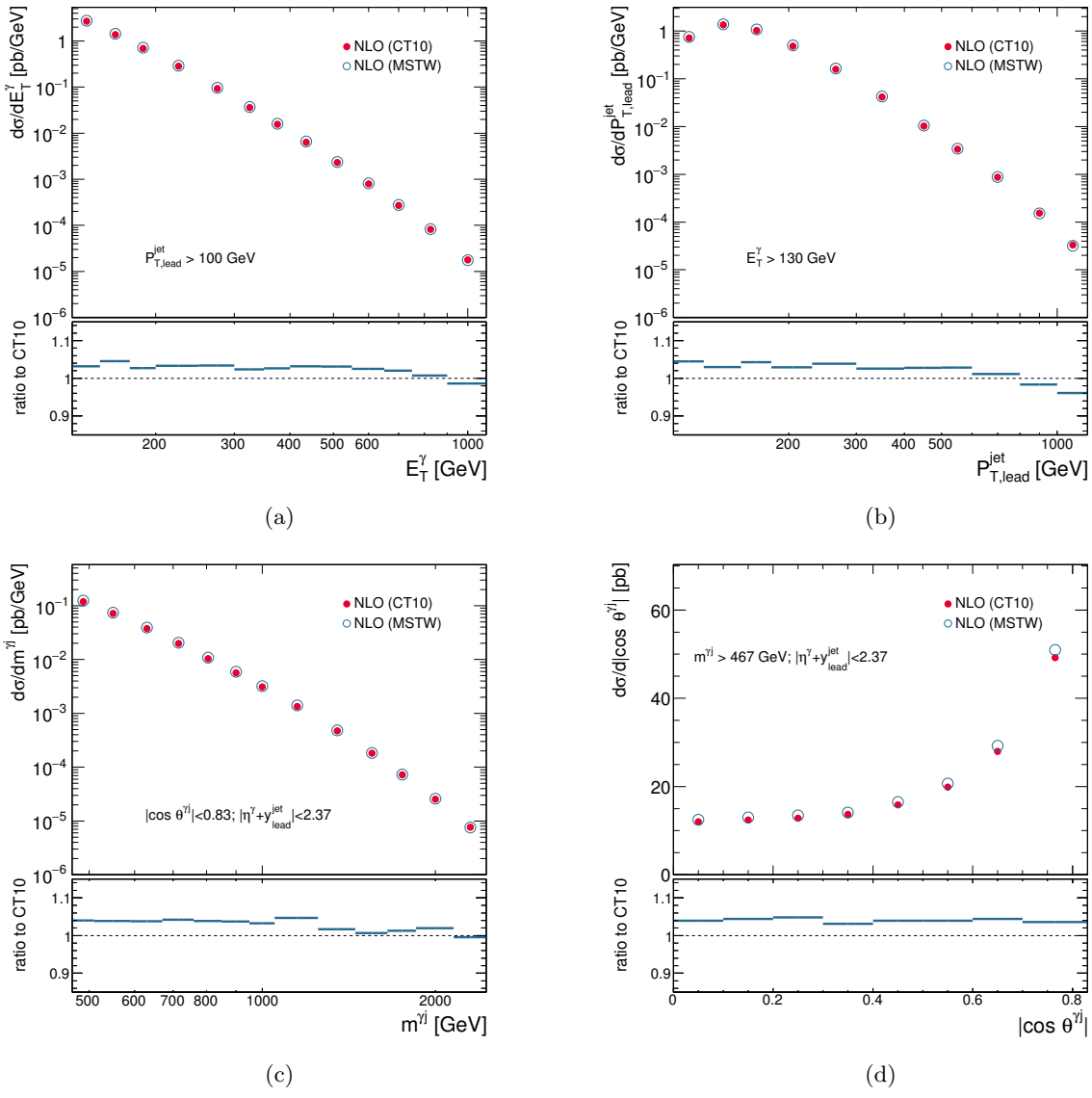


Figure 6.91: *Predicted NLO QCD cross sections using the CT10 (dots) and MSTW2008nlo (circles) parametrisations of the proton PDFs as functions of E^{γ}_{T} (a), $p^{jet}_{T,lead}$ (b), $m^{\gamma j}$ (c) and $|\cos\theta^{\gamma j}|$ (d) for the photon+one-jet sample. The lower part of the figures shows the ratio of the two NLO calculations.*

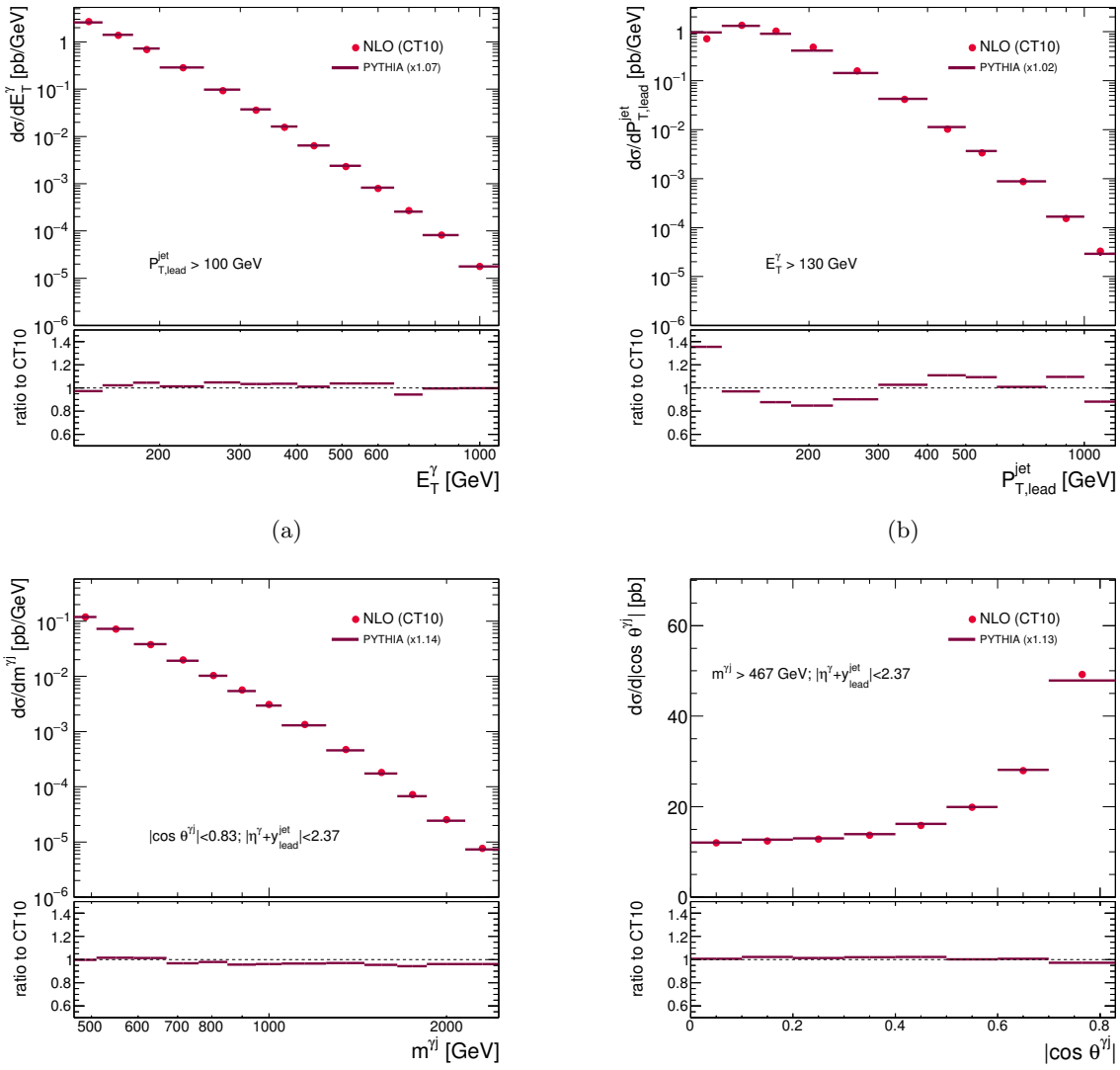


Figure 6.92: *Predicted NLO QCD cross sections using the CT10 (dots) parametrisations of the proton PDFs as functions of* E^{γ}_{T} *(a),* $p^{\mathrm{jet}}_{\mathrm{T,lead}}$ *(b),* $m^{\gamma j}$ *(c) and* $|\cos\theta^{\gamma j}|$ *(d) for the photon+one-jet sample. The predictions at parton level of* PYTHIA *(lines) are also shown. The lower part of the figures shows the ratio of the parton level MC and NLO QCD calculations.*

6.12.1 Hadronisation and underlying-event corrections to the NLO QCD calculations

The measurements performed in this analysis refer to jets of hadrons and include underlying-event (UE) effects, whereas the NLO QCD calculations refer to jets of partons without such an effect. The predictions were corrected to the hadron level with UE using the MC models. The correction factor, C_{NLO}, was defined as the ratio of the cross section for jets of hadrons with UE and that for jets of partons and was estimated by using PYTHIA.

This correction factor to the NLO QCD calculations was computed as

$$C_{\mathrm{NLO}} = C_{\mathrm{had}} \cdot C_{\mathrm{UE}}, \tag{6.13}$$

where C_{had} is the ratio of the cross sections for jets of hadrons and that for jets of partons; no UE effects were included in both cases. The C_{UE} correction factor is used to correct the NLO

QCD calculations to include the effects of the UE, so as to provide a fair comparison to the data. This correction factor was computed as the ratio of the cross section for jets of hadrons with UE and that for jets of hadrons without such an effect. Dedicated MC samples of PYTHIA generated without UE were used to obtain them.

For this method to be valid, the partonic level of the MC simulations must reasonably be close in shape to the NLO QCD predictions. The parton level predictions of PYTHIA give a reasonable description of the shape of the NLO QCD calculations (see figure 6.92).

Figure 6.93 shows the C_{NLO} correction factors as functions of the measured observables in the photon+one-jet sample. Figures 6.94 and 6.95 show the values of C_{had} and C_{UE}, respectively. The correction factors from PYTHIA are close to unity, except at high $p_{\mathrm{T,lead}}^{\mathrm{jet}}$; this effect comes from the UE effects. This effect in turn arises from the behaviour of this correction factor for the brem component in PYTHIA as can bee seen in figure 6.96 where the C_{UE} is shown separely for both components as a function of the measured observables in the photon+one-jet sample.

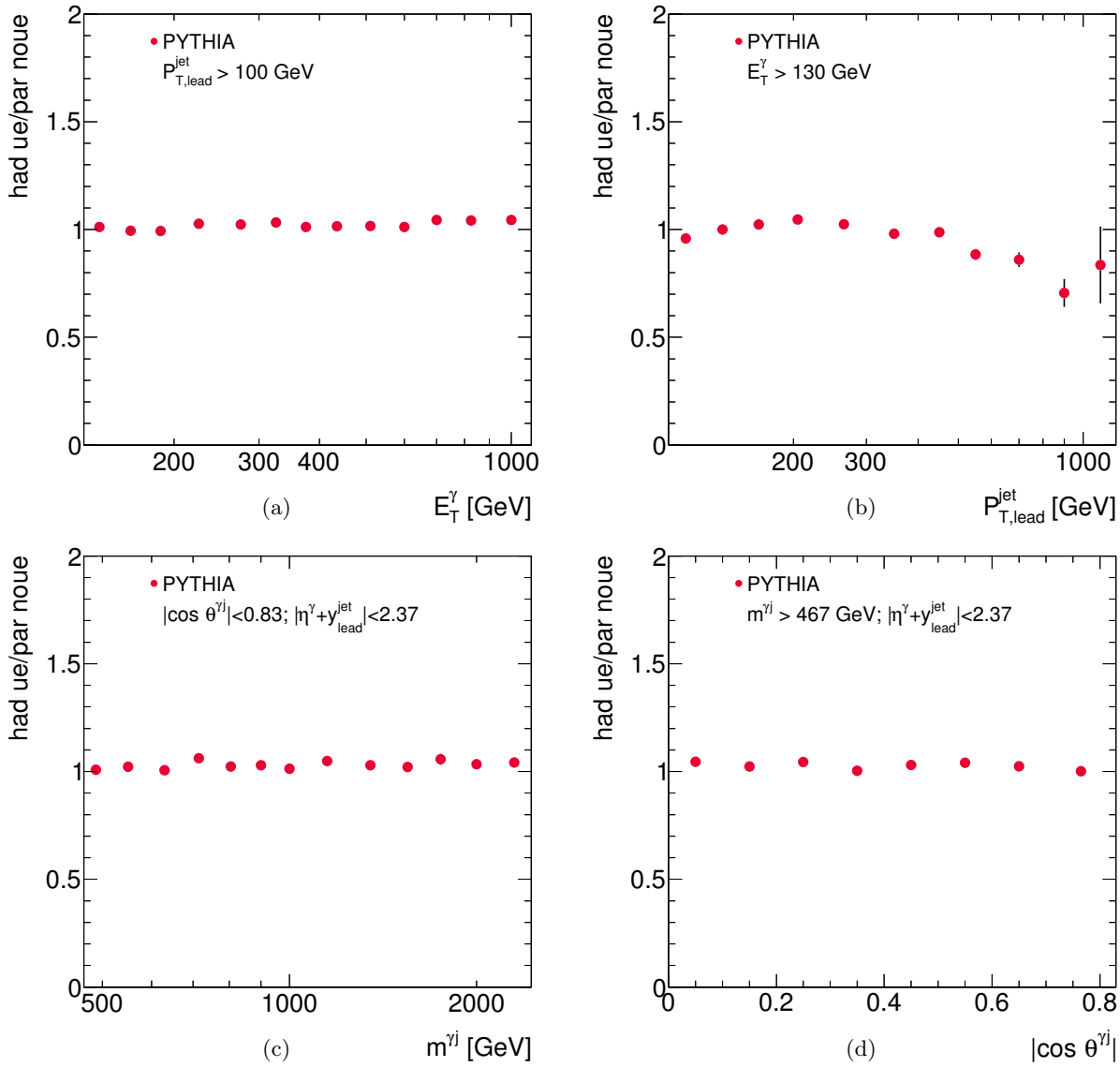


Figure 6.93: *C_{NLO} correction factors as functions of E_{T}^{γ} (a), $p_{\mathrm{T,lead}}^{\mathrm{jet}}$ (b), $m^{\gamma j}$ (c) and $|\cos\theta^{\gamma j}|$ (d) for the photon+one-jet sample.*

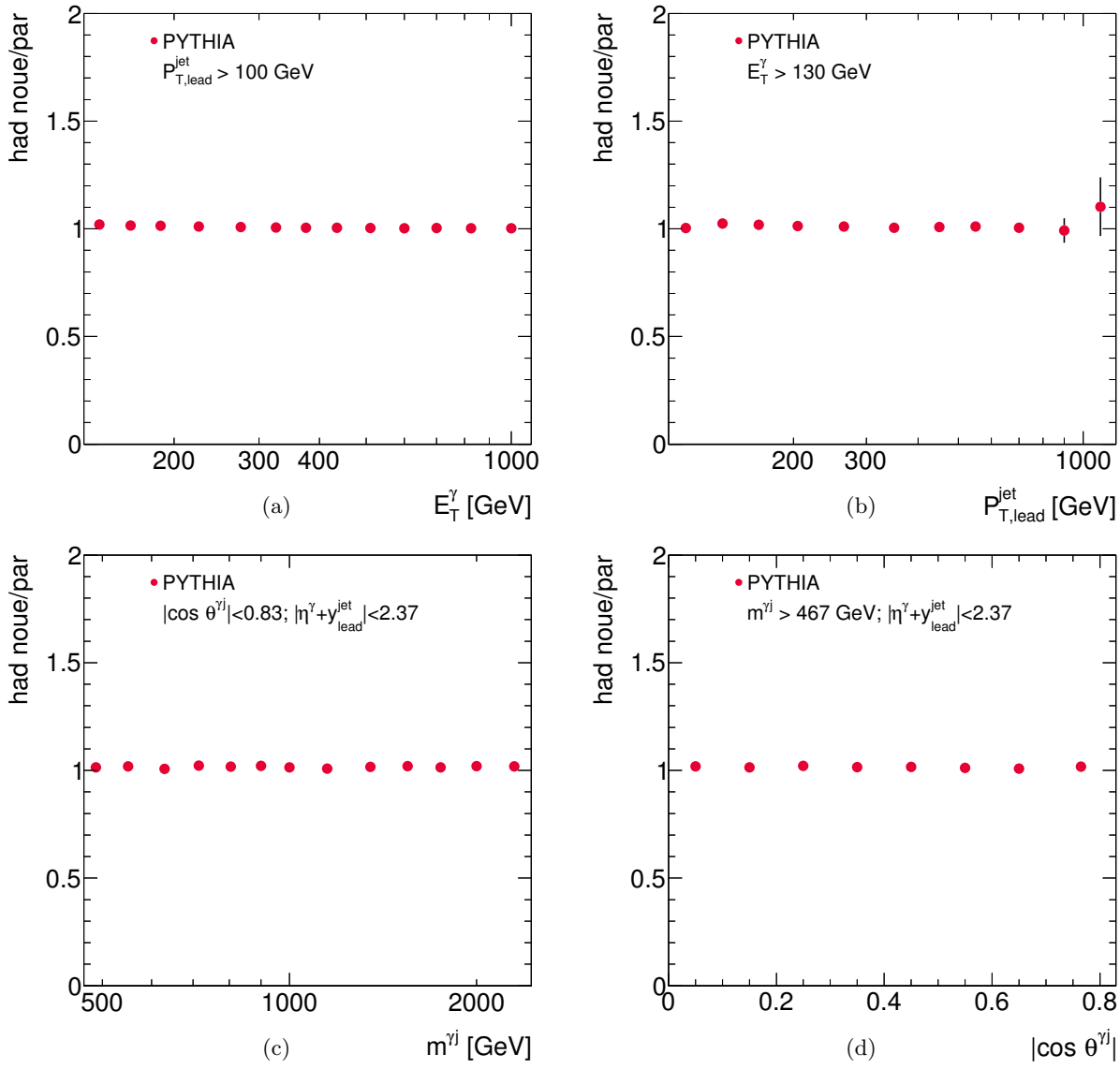


Figure 6.94: *C_{had} correction factors as functions of E_{T}^{γ} (a), $p_{\mathrm{T,lead}}^{\mathrm{jet}}$ (b), $m^{\gamma j}$ (c) and $|\cos\theta^{\gamma j}|$ (d) for the photon+one-jet sample.*

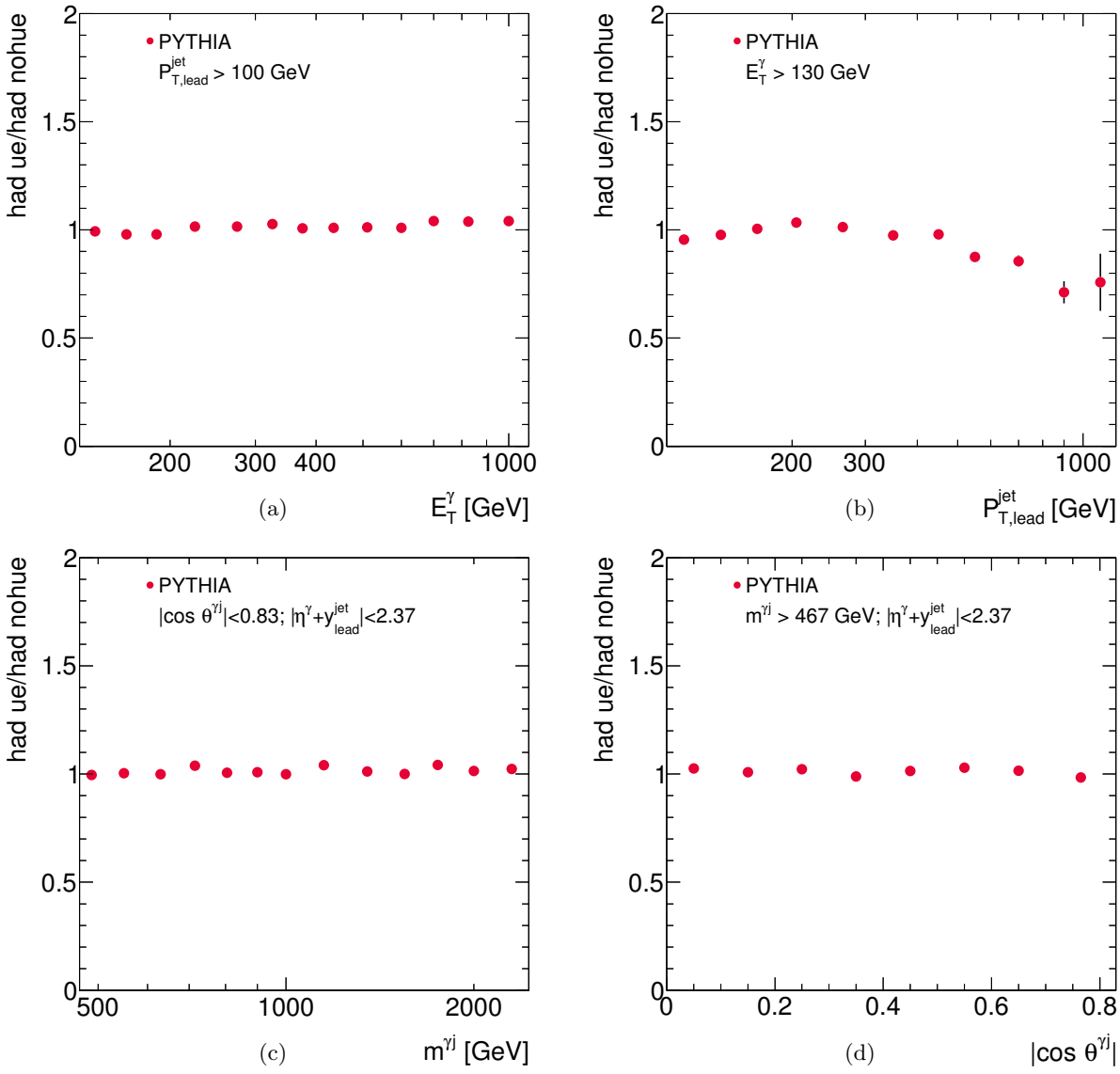


Figure 6.95: *C_{UE} correction factors as functions of E_{T}^{γ} (a), $p_{\mathrm{T,lead}}^{\mathrm{jet}}$ (b), $m^{\gamma j}$ (c) and $|\cos\theta^{\gamma j}|$ (d) for the photon+one-jet sample.*

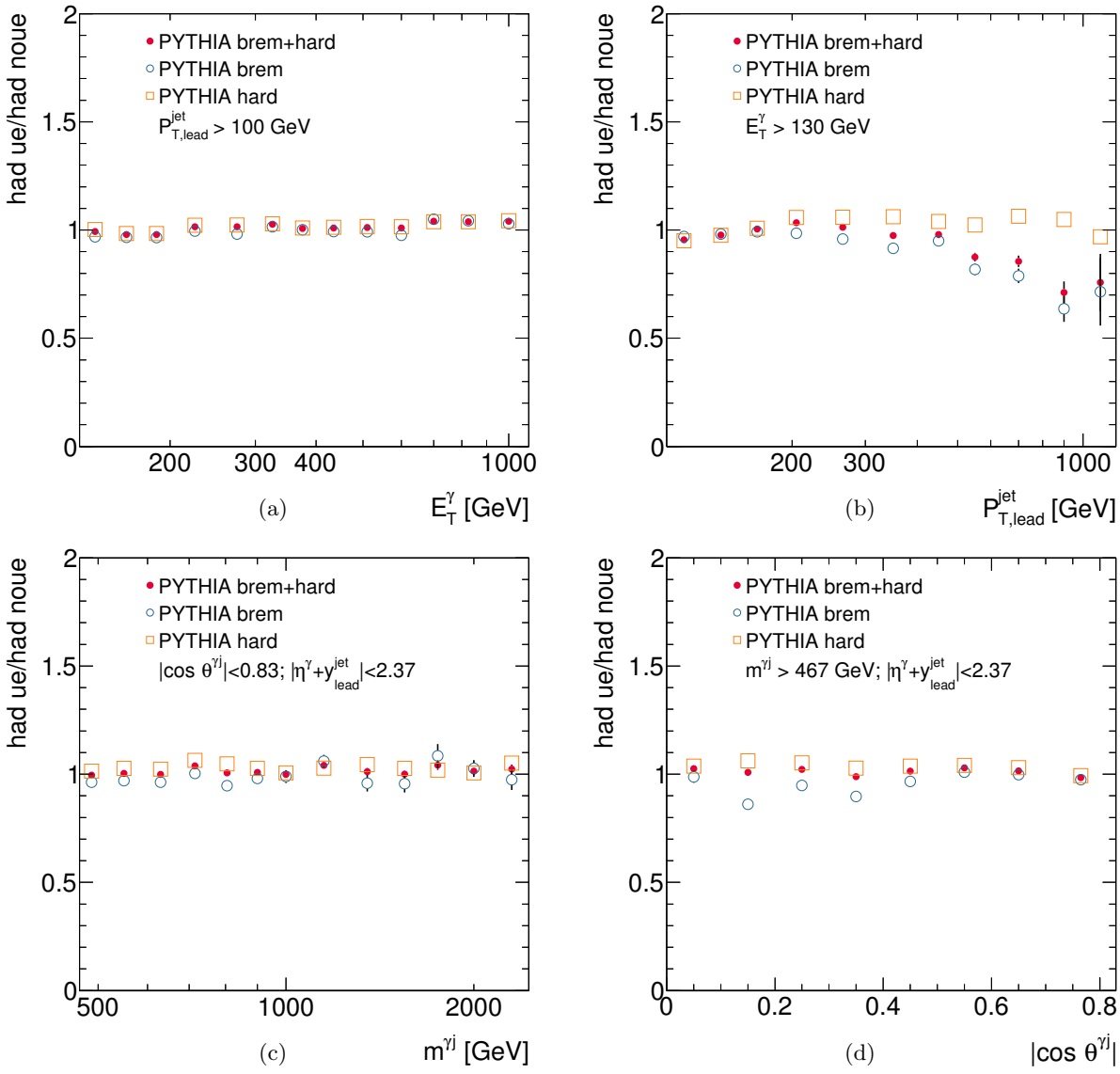


Figure 6.96: C_{UE} *correction factors from* PYTHIA *(dots),* PYTHIA *brem (open circles) and* PYTHIA *hard (open squares) as functions of* E_{T}^{γ} *(a),* $p_{\mathrm{T,lead}}^{\mathrm{jet}}$ *(b),* $m^{\gamma j}$ *(c) and* $|\cos\theta^{\gamma j}|$ *(d) for the photon+one-jet sample.*

6.12.2 Theoretical uncertainties

The following sources of uncertainty in the theoretical NLO predictions were considered:

- the uncertainty on the NLO QCD calculations due to the terms beyond NLO was estimated by repeating the calculations using values of μ_R, μ_F and μ_f scaled by factors 0.5 and 2. The three scales were varied individually;
- the uncertainty on the NLO QCD calculations due to those on the parametrisation of the proton PDFs was estimated by repeating the calculations using the 52 additional sets from the CT10 error analysis;
- the uncertainty on the NLO QCD calculations due to that on the value of $\alpha_s(m_Z)$ was estimated by repeating the calculations using two additional sets of parametrisations of the proton PDFs, for which different values of $\alpha_s(m_Z)$ were assumed in the fits, namely $\alpha_s(m_Z) = 0.117$ and $\alpha_s(m_Z) = 0.119$;

The dominant theoretical uncertainty is that arising from the terms beyond NLO. It was checked [90] that the effect of the simultaneous variation of the scales (in pairs of the three scales simultaneously) is not larger than that arising fom the individual variations of μ_R; therefore, only the individual variations of the scales have been included as theoretical uncertainty.

Figures 6.97 to figures 6.101 show the effect of the relative theoretical uncertainties in the kinematic region of the measurements as functions of each of the observables measured in the photon+one-jet sample. The total theoretical uncertainty was obtained by adding in quadrature the individual uncertainties listed above. Figure 6.102 shows the total theoretical uncertainties as functions of each of the observables measured in the photon+one-jet sample.

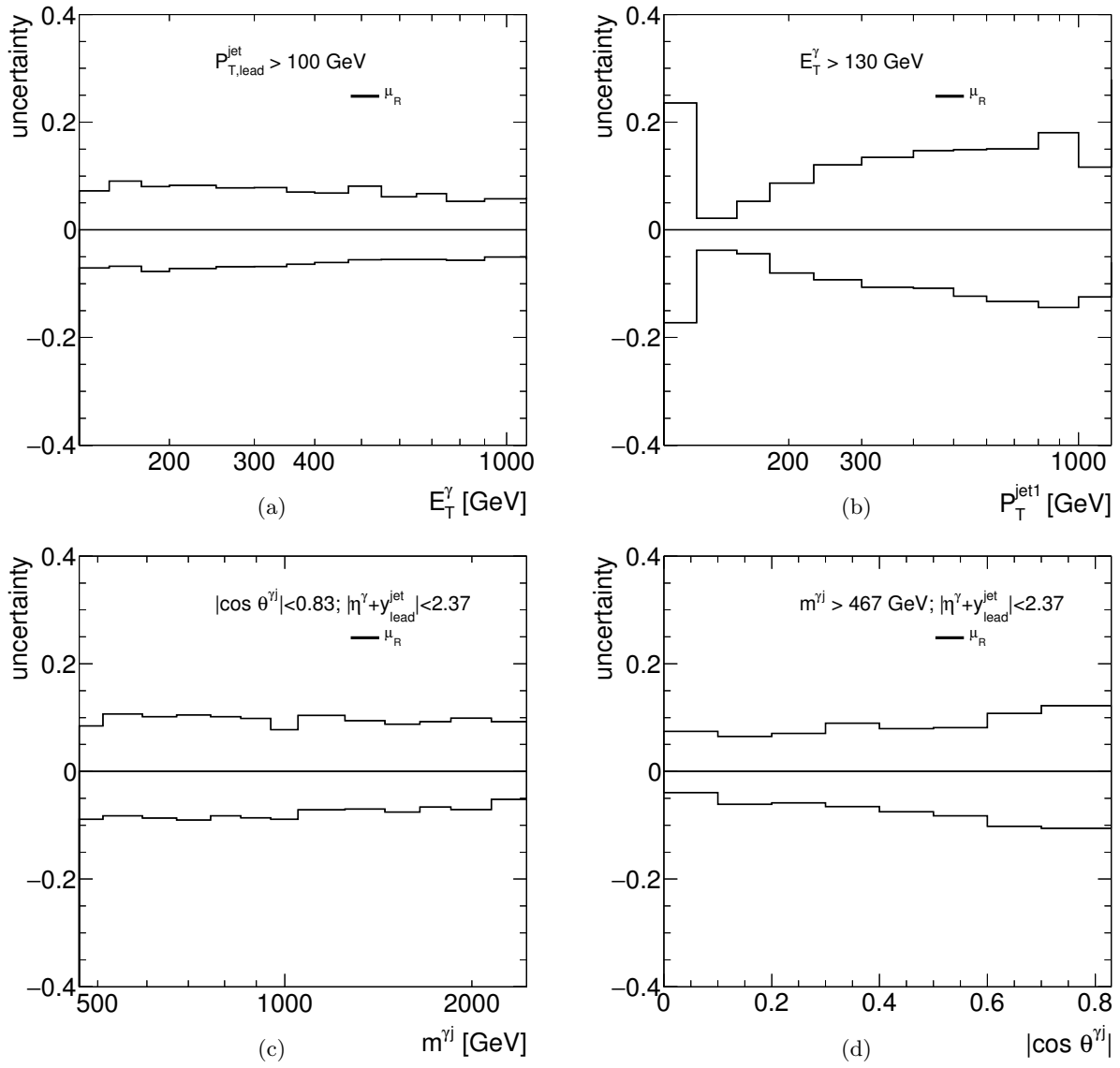


Figure 6.97: *Theoretical uncertainties arising from terms beyond NLO (variation of μ_R) as functions of E^{γ}_{T} (a), $p^{\mathrm{jet}}_{\mathrm{T,lead}}$ (b), $m^{\gamma j}$ (c) and $|\cos\theta^{\gamma j}|$ (d) for the photon+one-jet sample.*

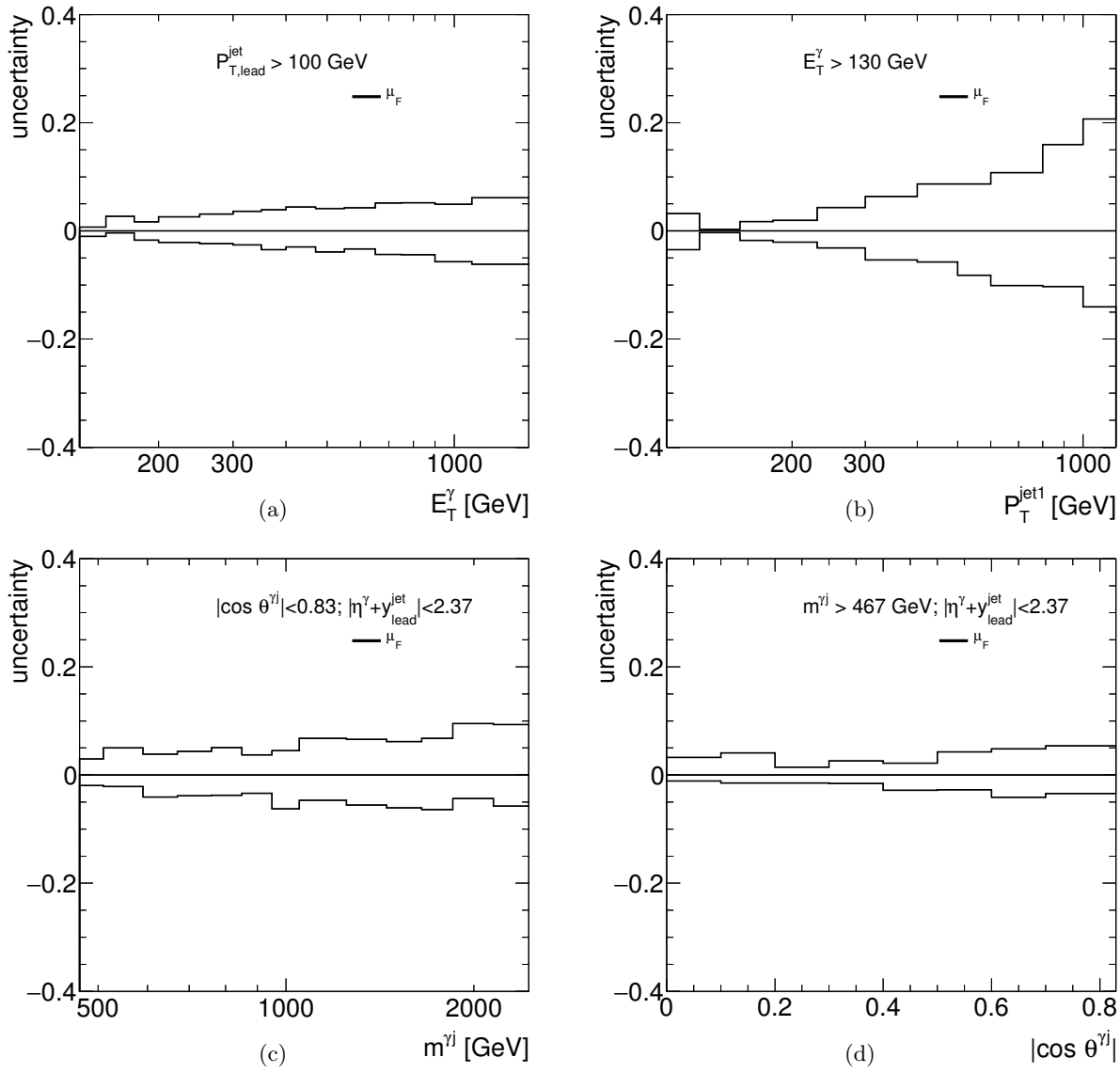


Figure 6.98: *Theoretical uncertainties arising from terms beyond NLO (variation of μ_F) as functions of E_{T}^{γ} (a), $p_{\mathrm{T,lead}}^{\mathrm{jet}}$ (b), $m^{\gamma j}$ (c) and $|\cos\theta^{\gamma j}|$ (d) for the photon+one-jet sample.*

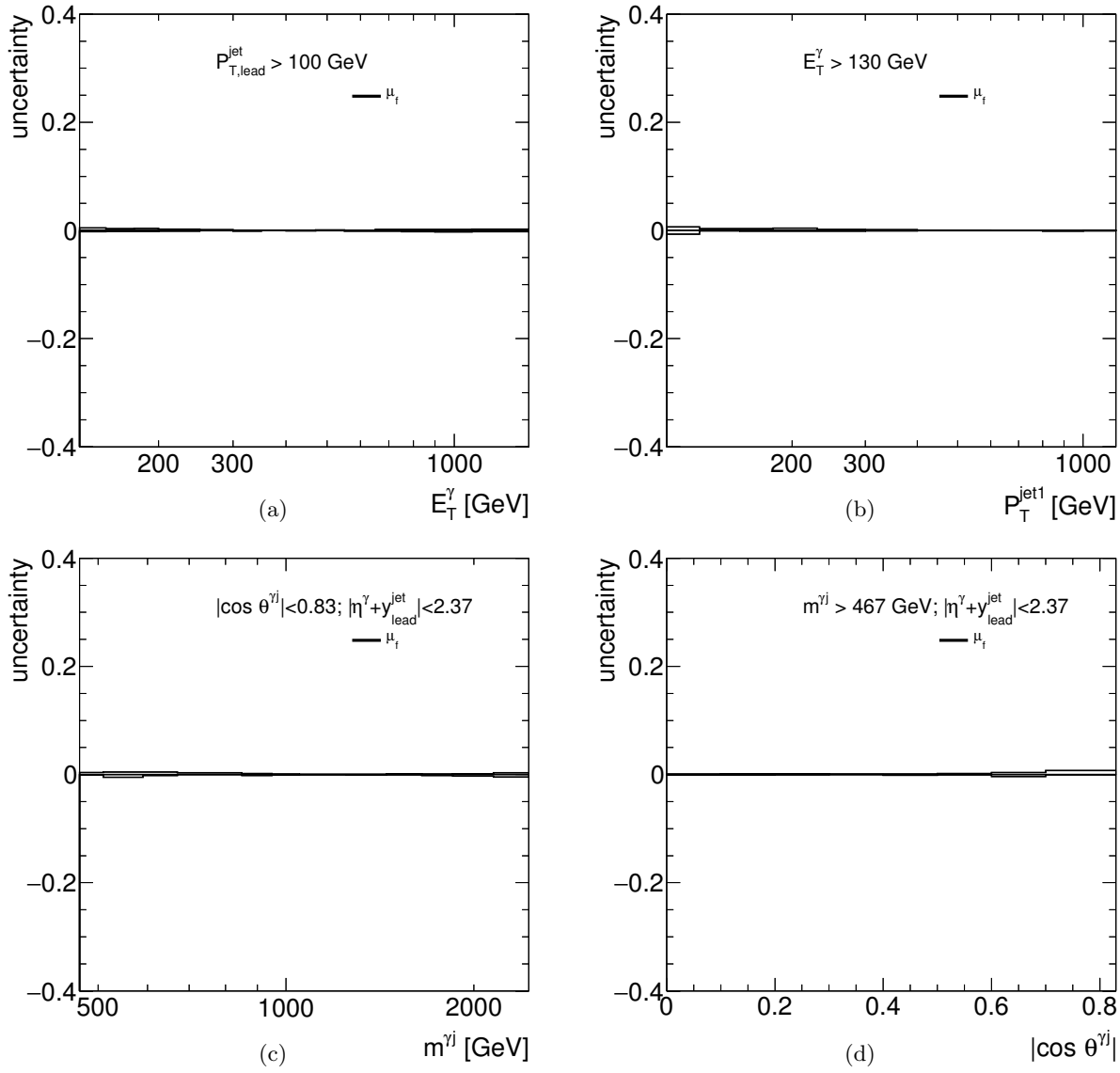


Figure 6.99: *Theoretical uncertainties arising from terms beyond NLO (variation of μ_f) as functions of E_{T}^{γ} (a), $p_{\mathrm{T,lead}}^{\mathrm{jet}}$ (b), $m^{\gamma j}$ (c) and $|\cos\theta^{\gamma j}|$ (d) for the photon+one-jet sample.*

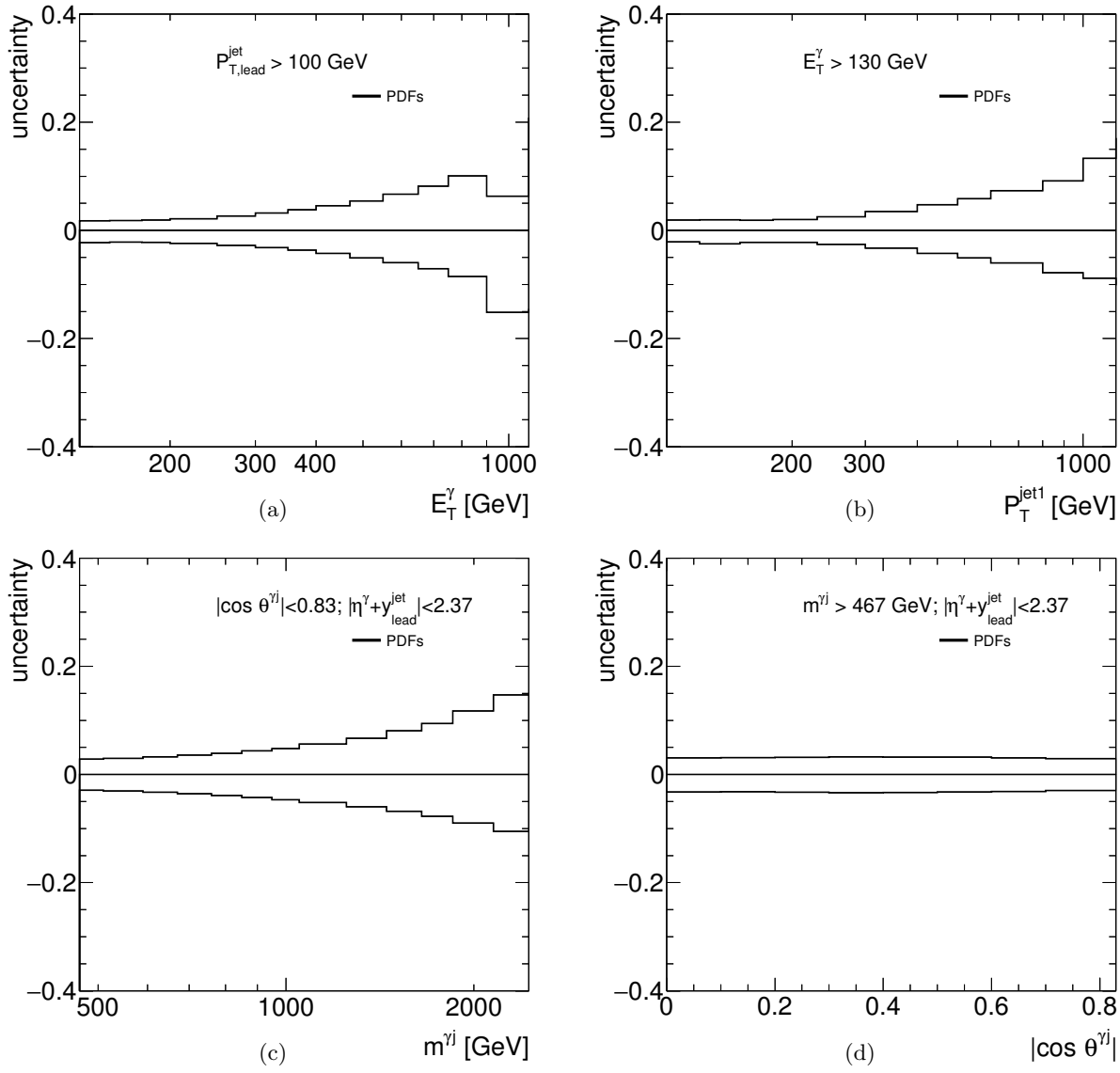


Figure 6.100: *Theoretical uncertainties arising from the uncertainty in the PDFs as functions of* E^{γ}_{T} *(a),* $p^{\mathrm{jet}}_{\mathrm{T,lead}}$ *(b),* $m^{\gamma j}$ *(c) and* $|\cos\theta^{\gamma j}|$ *(d) for the photon+one-jet sample.*

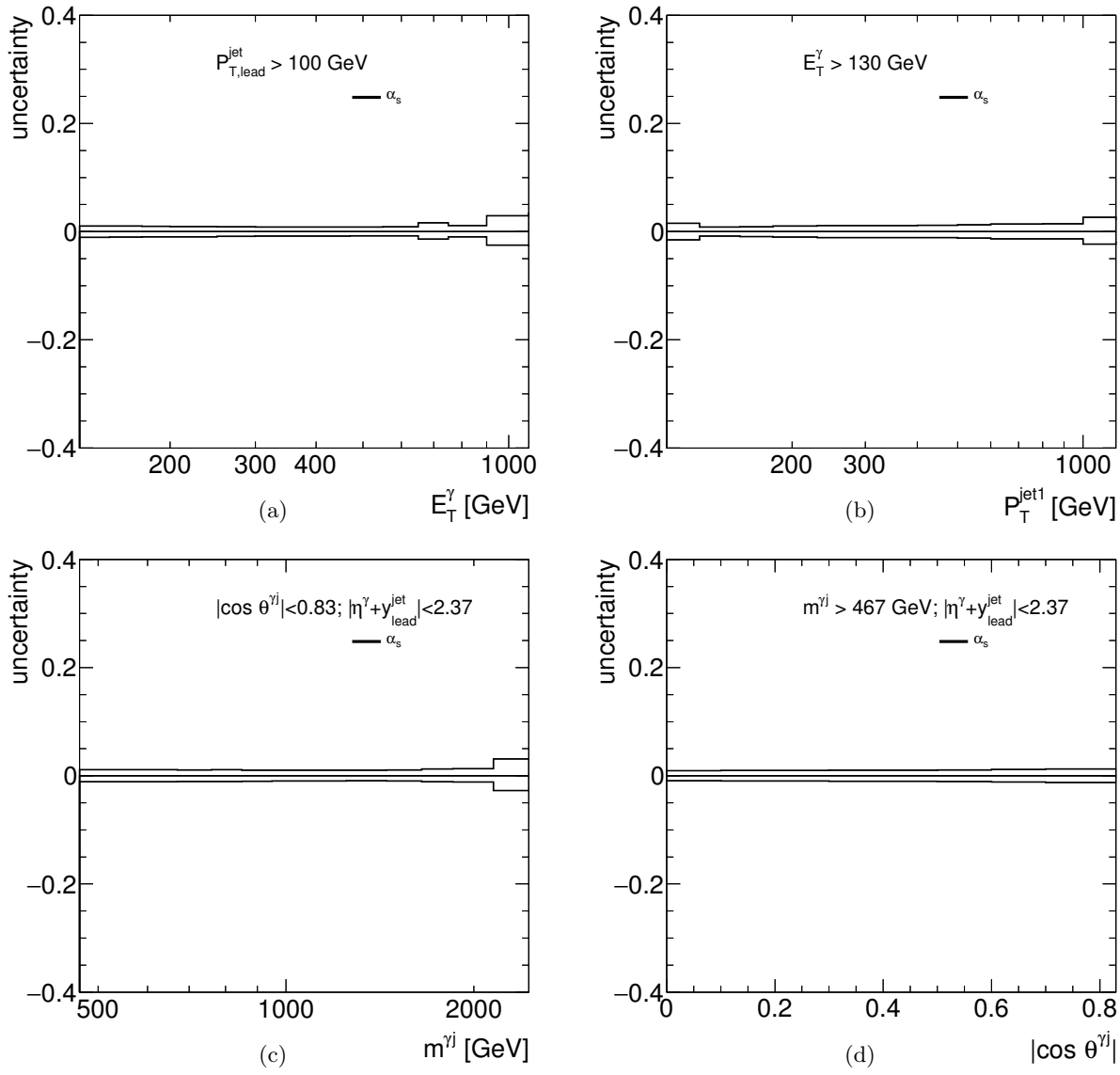


Figure 6.101: *Theoretical uncertainties arising from the uncertainty in α_s as functions of E^{γ}_{T} (a), $p^{\mathrm{jet}}_{\mathrm{T,lead}}$ (b), $m^{\gamma j}$ (c) and $|\cos\theta^{\gamma j}|$ (d) for the photon+one-jet sample.*

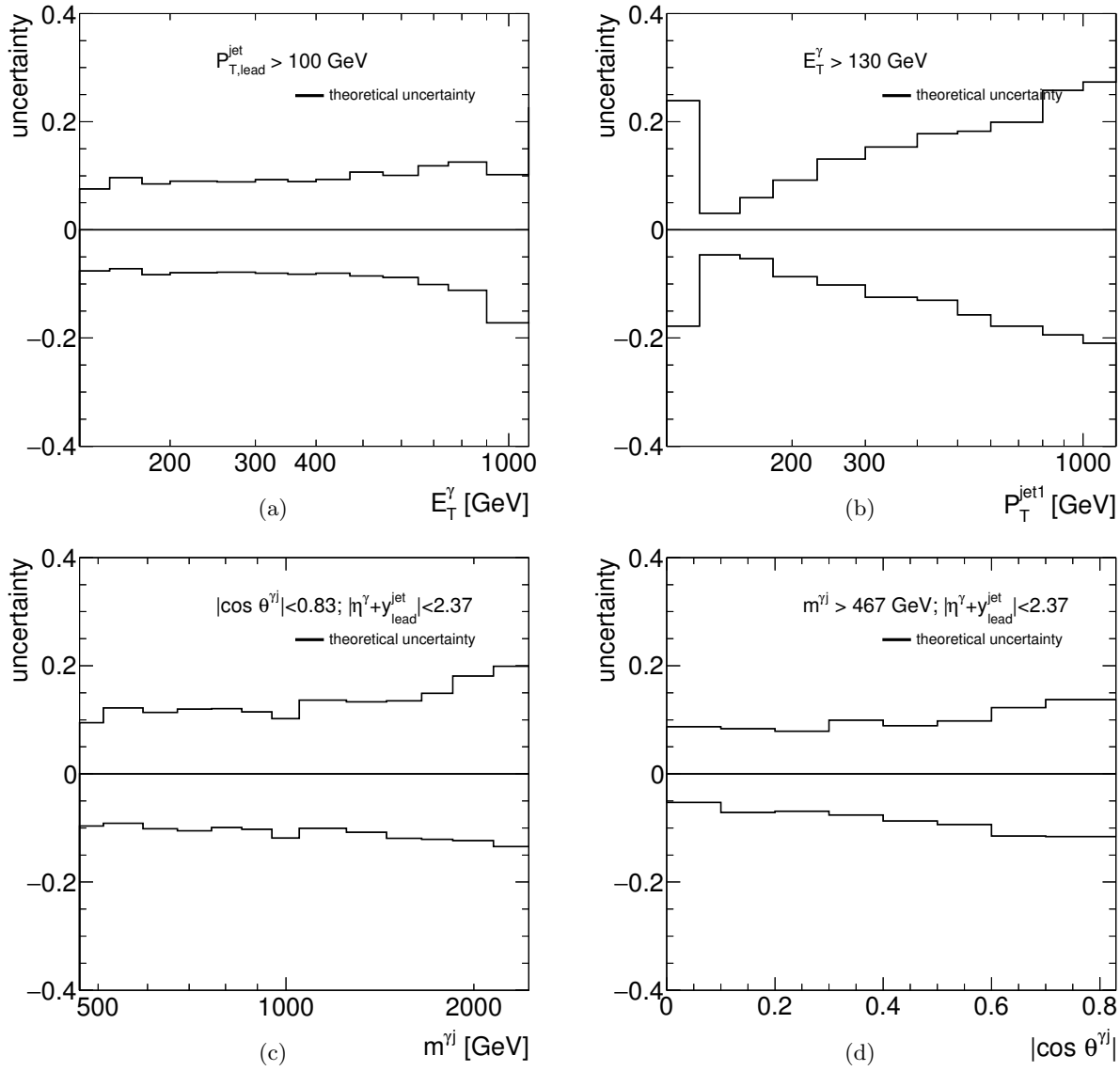


Figure 6.102: *Total theoretical uncertainties as functions of* E_{T}^{γ} *(a),* $p_{\mathrm{T,lead}}^{\mathrm{jet}}$ *(b),* $m^{\gamma j}$ *(c) and* $|\cos\theta^{\gamma j}|$ *(d) for the photon+one-jet sample.*

6.13 Results

In this section the results obtained for the different photon+jets samples (see section 6.4) are shown and discussed.

6.13.1 Isolated photon plus one jet production

Figure 6.103 shows the isolated photon plus one jet differential cross sections as functions of E_{T}^{γ}, $p_{\mathrm{T,lead}}^{\mathrm{jet}}$, $m^{\gamma j}$ and $|\cos\theta^{\gamma j}|$.

The measured $d\sigma/dE_{\mathrm{T}}^{\gamma}$ (figure 6.103a) decreases by almost five orders of magnitude as E_{T}^{γ} increases. Values up to $E_{\mathrm{T}}^{\gamma} = 1.1$ TeV are measured. The NLO QCD calculation provided by JETPHOX gives a good description of the data within the experimental and theoretical uncertainties. The experimental uncertainty is below 6% for $E_{\mathrm{T}}^{\gamma} < 650$ GeV, dominated by the photon ES; for $E_{\mathrm{T}}^{\gamma} > 650$ GeV, it grows up to 15% at $E_{\mathrm{T}}^{\gamma} = 1$ TeV and is dominated by the statistical uncertainty in this region. The theoretical uncertainty is rather constant and $\approx 10\%$; it is dominated by the contribution arising from the higher order terms, in particular, the one arising from the variation of μ_R, though for $E_{\mathrm{T}}^{\gamma} > 500$ GeV the uncertainty due to that from the PDFs grows to be of the same order and dominates for higher E_{T}^{γ} values.

The measured $d\sigma/dp_{\mathrm{T,lead}}^{\mathrm{jet}}$ (figure 6.103b) decreases by five orders of magnitude from $p_{\mathrm{T,lead}}^{\mathrm{jet}} > 120$ GeV up to the highest measured value, $p_{\mathrm{T,lead}}^{\mathrm{jet}} = 1.2$ TeV; for $p_{\mathrm{T,lead}}^{\mathrm{jet}} < 120$ GeV, the cross section decreases due to the kinematic constraints. The NLO QCD calculations give a good description of the data, except for $p_{\mathrm{T,lead}}^{\mathrm{jet}} < 120$ GeV, where the calculation includes only the lowest non-trivial order due to the asymmetric cut between the photon and the leading jet, i.e. the leading order contribution is zero. The theoretical uncertainty in this region is large precisely for the same reason, since the uncertainty arising from the higher order terms is also bigger. The experimental uncertainty is $\approx 7\%$, except in the last bin ($\approx 15\%$, dominated by the statistical uncertainty) and is dominated by the JES. The theoretical uncertainty, for $p_{\mathrm{T,lead}}^{\mathrm{jet}} > 120$ GeV, increases from $\approx 5\%$ to $\approx 25\%$ for $p_{\mathrm{T,lead}}^{\mathrm{jet}} = 1.2$ TeV and is dominated by the uncertainty arising from the variation of μ_R.

The measured $d\sigma/dm^{\gamma j}$ (figure 6.103c) decreases by four orders of magnitude up to the highest measured value $m^{\gamma j} = 2.45$ TeV. The NLO QCD calculations give a good description of the data and no significant deviation with respect to the QCD predictions is observed down to scales of $\hbar/m^{\gamma j} \approx 8 \cdot 10^{-5}$ fm. The experimental uncertainty for $m^{\gamma j} < 1.5$ TeV is below 5% and it grows up to 20% in the last bin. The dominant uncertainties are the photon ES, the JES and the model dependence; for $m^{\gamma j} > 1.5$ TeV, the statistical uncertainty dominates. The theoretical uncertainty is $\approx 10\,(20)\%$ at $m^{\gamma j} = 467\,(2450)$ GeV and is dominated by the contribution arising from higher-order terms, in particular the one coming from the variation of μ_R ($\approx 10\%$), though for $m^{\gamma j} > 1.65$ TeV the uncertainty due to that from the PDFs grows to be of the same order and dominates for higher $m^{\gamma j}$ values.

The measured $d\sigma/d|\cos\theta^{\gamma j}|$ (figure 6.103d) increases as $|\cos\theta^{\gamma j}|$ increases. The NLO QCD calculations give a good description of the data. The experimental uncertainty is $\approx 4\%$, dominated by the contributions from the photon ES, the JES and the model dependence. The theoretical uncertainty is $\approx 10\%$, dominated by the contribution arising from the higher-order terms, in particular the one coming from the variation of μ_R.

All these comparison provide stringent tests of pQCD and validate the description of the dynamics of isolated-photon plus one jet production in pp collisions at $\mathcal{O}(\alpha_{em}\alpha_s^2)$

To gain further insight into the dynamics of the photon-jet system, differential cross sections as functions of $|\cos\theta^{\gamma j}|$ were measured in different regions of $m^{\gamma j}$. Figure 6.104 shows the measured and predicted normalised differential cross sections and figure 6.105 shows their ratio.

In these figures, all the cross sections were normalised to unity at $|\cos\theta^{\gamma j}| = 0.1$ so that the increase of the cross section at $|\cos\theta^{\gamma j}| = 0.8$ can be clearly observed as $m^{\gamma j}$ increases. The NLO calculations describe well the scale evolution of the measured differential cross sections.

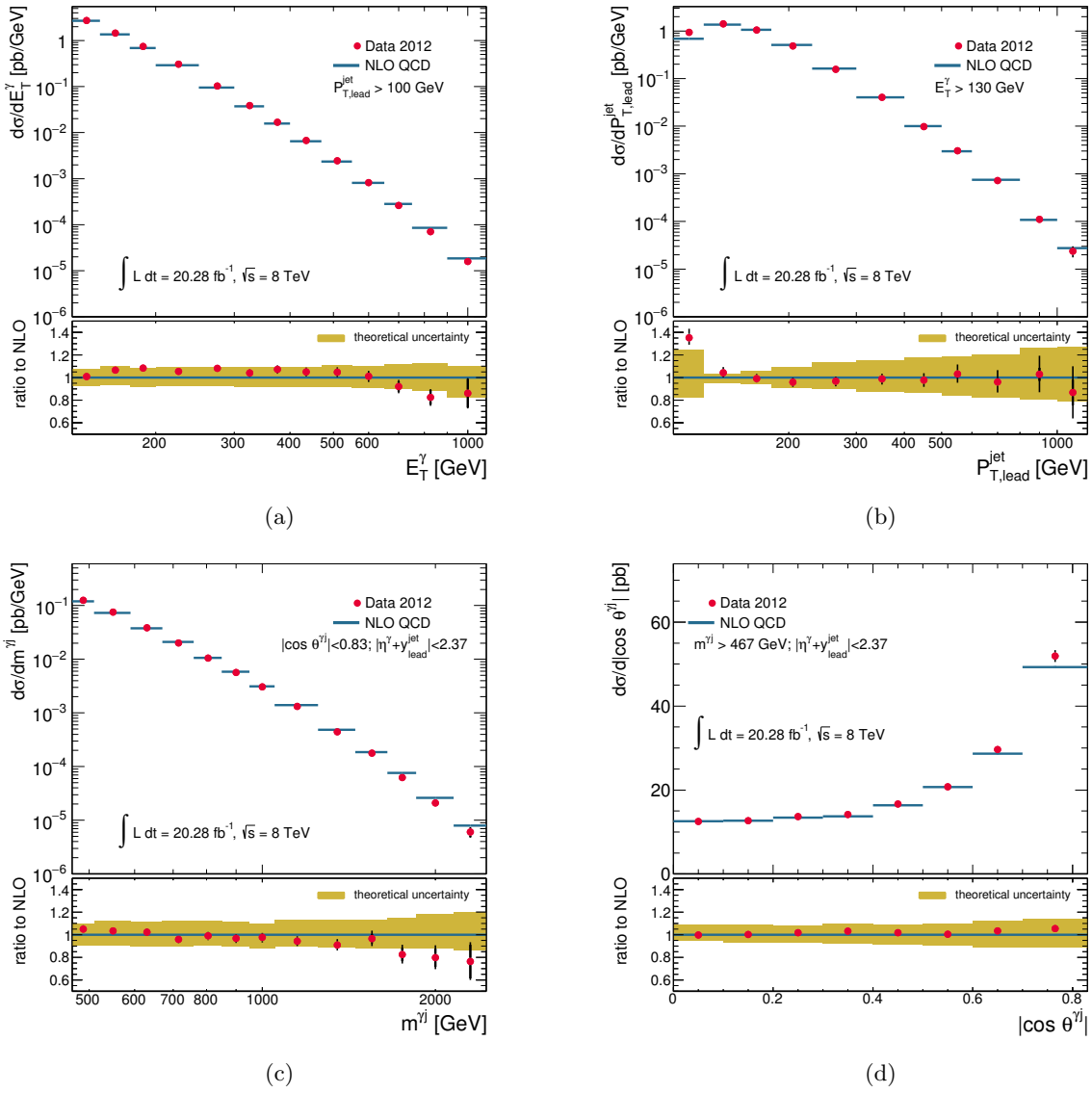


Figure 6.103: *Measured differential cross sections for isolated photon plus one jet production (dots) as functions of E^{γ}_{T} (a), $p^{\mathrm{jet}}_{\mathrm{T,lead}}$ (b), $m^{\gamma j}$ (c) and $|\cos\theta^{\gamma j}|$ (d). The NLO QCD calculations from* JETPHOX *corrected for hadronisation and underlying-event effects and using the CT10 parametrisation for the proton PDFs are also shown. The bottom part of each figure shows the ratio of the measured cross section and the NLO QCD calculation. The inner (outer) error bars represent the statistical uncertainties (the statistical and systematic uncertainties added in quadrature) and the shaded band represents the theoretical uncertainty. For most of the points, the inner error bars are smaller than the marker size and, thus, not visible.*

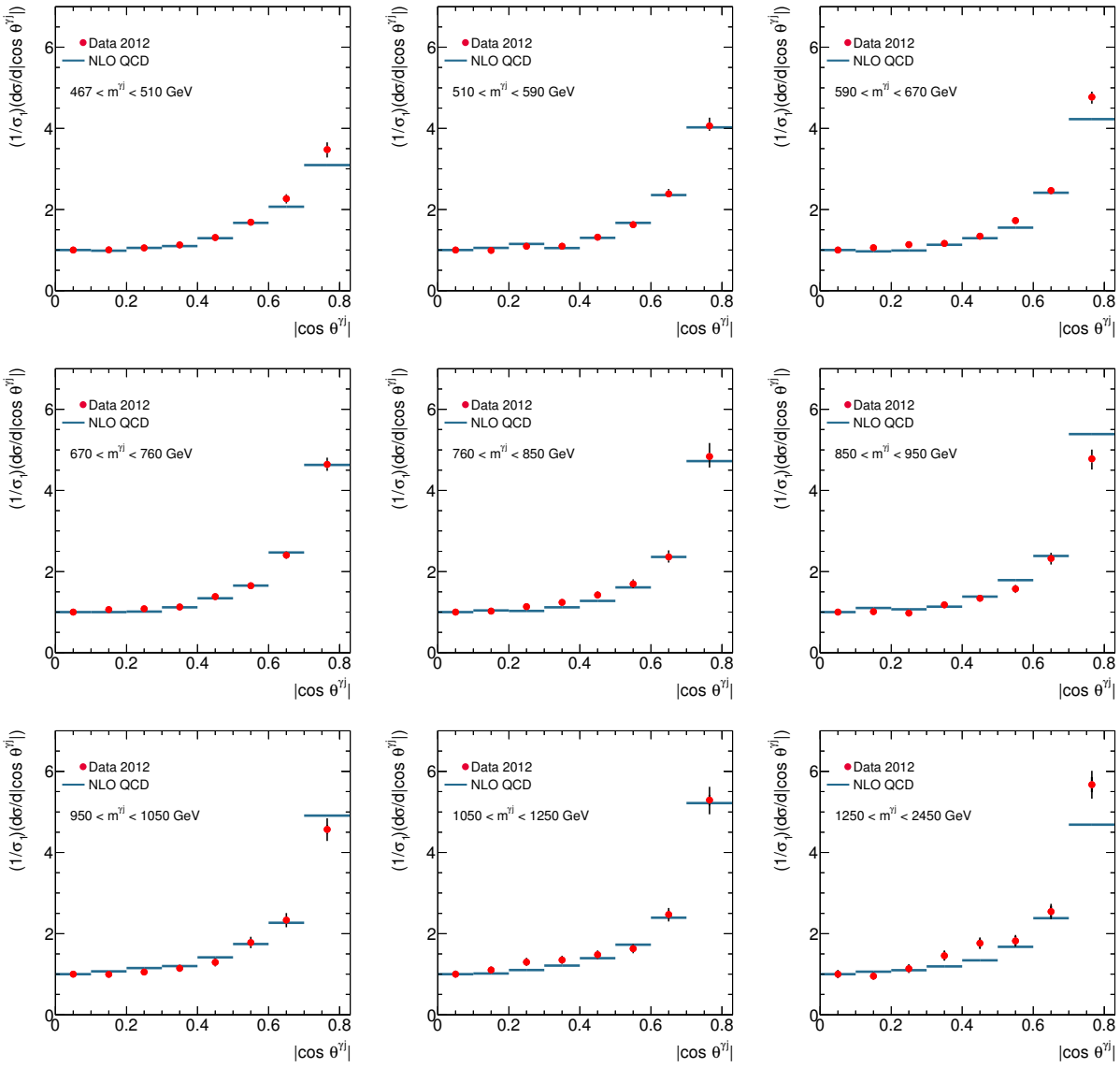


Figure 6.104: *Measured normalised differential cross sections for isolated photon plus one jet production (dots) as functions of* $|\cos\theta^{\gamma j}|$ *in different regions of* $m^{\gamma j}$*. Other details as in the caption to figure 6.103.*

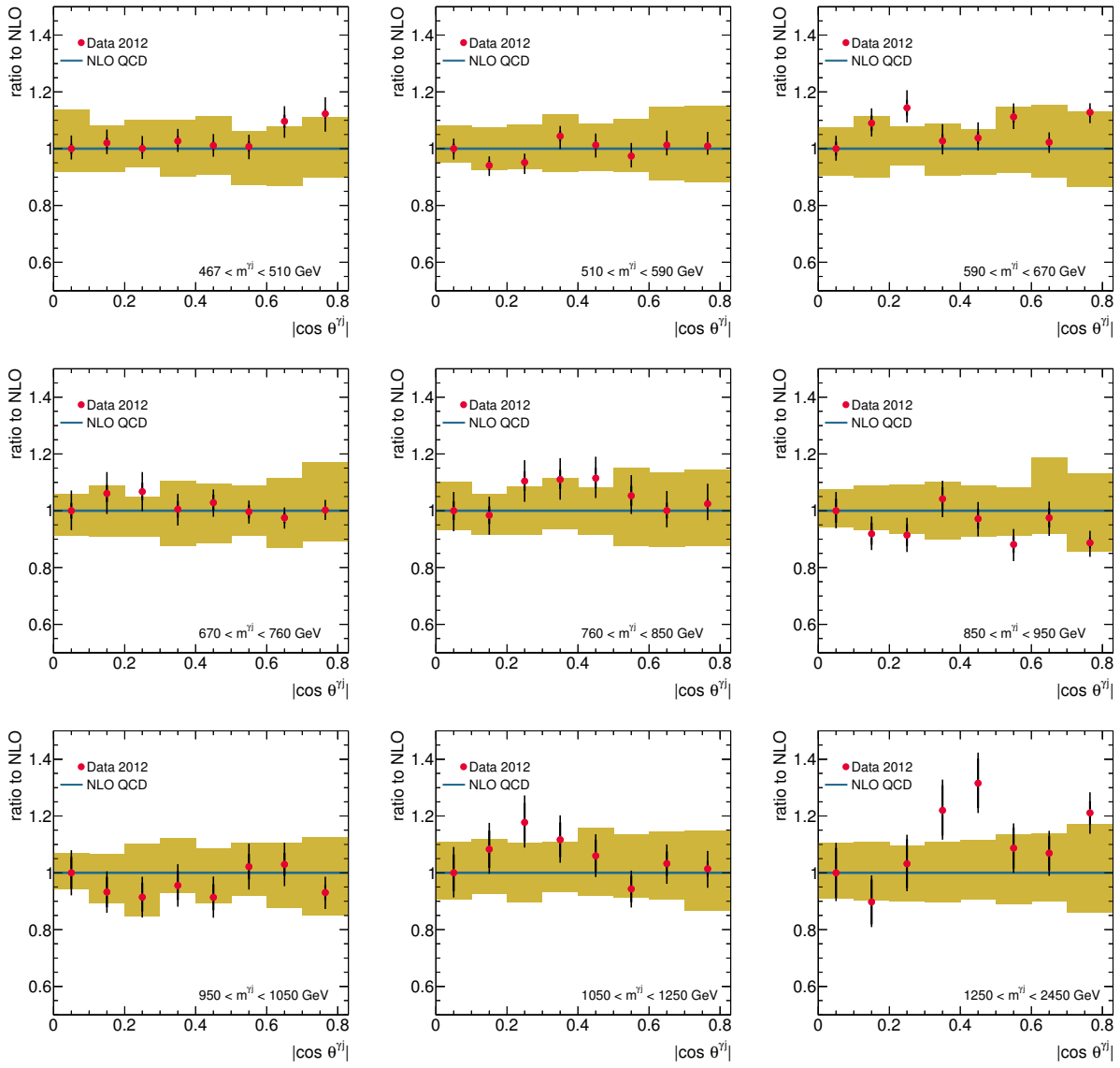


Figure 6.105: *Ratio of the measured normalised differential cross sections for isolated photon plus one jet production and the NLO QCD calculation (dots) as functions of* $|\cos\theta^{\gamma j}|$ *in different regions of* $m^{\gamma j}$*. Other details as in the caption to figure 6.103.*

6.13.2 Isolated photon plus two jets production

Figure 6.106 shows the isolated photon plus two jet differential cross sections as functions of E_{T}^{γ}, $p_{\mathrm{T,sublead}}^{\mathrm{jet}}$, $\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$ and $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$. The measured $d\sigma/dE_{\mathrm{T}}^{\gamma}$ and $d\sigma/dp_{\mathrm{T,sublead}}^{\mathrm{jet}}$ decrease by five orders of magnitude within the measured range. The LO matrix-element plus parton shower prediction of Pythia gives a good description of $d\sigma/dE_{\mathrm{T}}^{\gamma}$ whereas the prediction of Sherpa describes well the measured $d\sigma/dp_{\mathrm{T,sublead}}^{\mathrm{jet}}$.

The measured $d\sigma/d\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$ and $d\sigma/d\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$ display a maximum at approximately 2-2.5. The predictions of Sherpa give a good description of the data, whereas the predictions of Pythia have a very different shape than the data. This can be explained by the fact that in the $2 \to 2$ predictions of Pythia, a second jet can arise only from the parton shower, whereas in Sherpa, $2 \to 3$ and $2 \to 4$ matrix elements contributions can provide a second hard jet.

The scale evolution of the photon plus two-jet production was tested by measuring the azimuthal angle differences for $E_{\mathrm{T}}^{\gamma} < 300$ GeV and $E_{\mathrm{T}}^{\gamma} > 300$ GeV. Figure 6.107 shows the $\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$ and $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$ normalised differential cross sections for the two E_{T}^{γ} ranges; in this figure, the cross sections were normalised to unity at $\Delta\phi = \pi/10$ so that the change in shape at large $\Delta\phi$ values for the two E_{T}^{γ} ranges can be clearly observed. For $E_{\mathrm{T}}^{\gamma} < 300$ GeV, both cross sections have similar shape to the measured in the $E_{\mathrm{T}}^{\gamma} > 130$ GeV; however, for $E_{\mathrm{T}}^{\gamma} > 300$ GeV, the $\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$ cross section is enhanced for $\Delta\phi^{\gamma-\mathrm{jet\,sublead}} > 2$ and the $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$ cross section is suppressed for $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}} > 1$. The predictions of Sherpa describe well the shape of the measured normalised differential cross sections and their evolution with the scale.

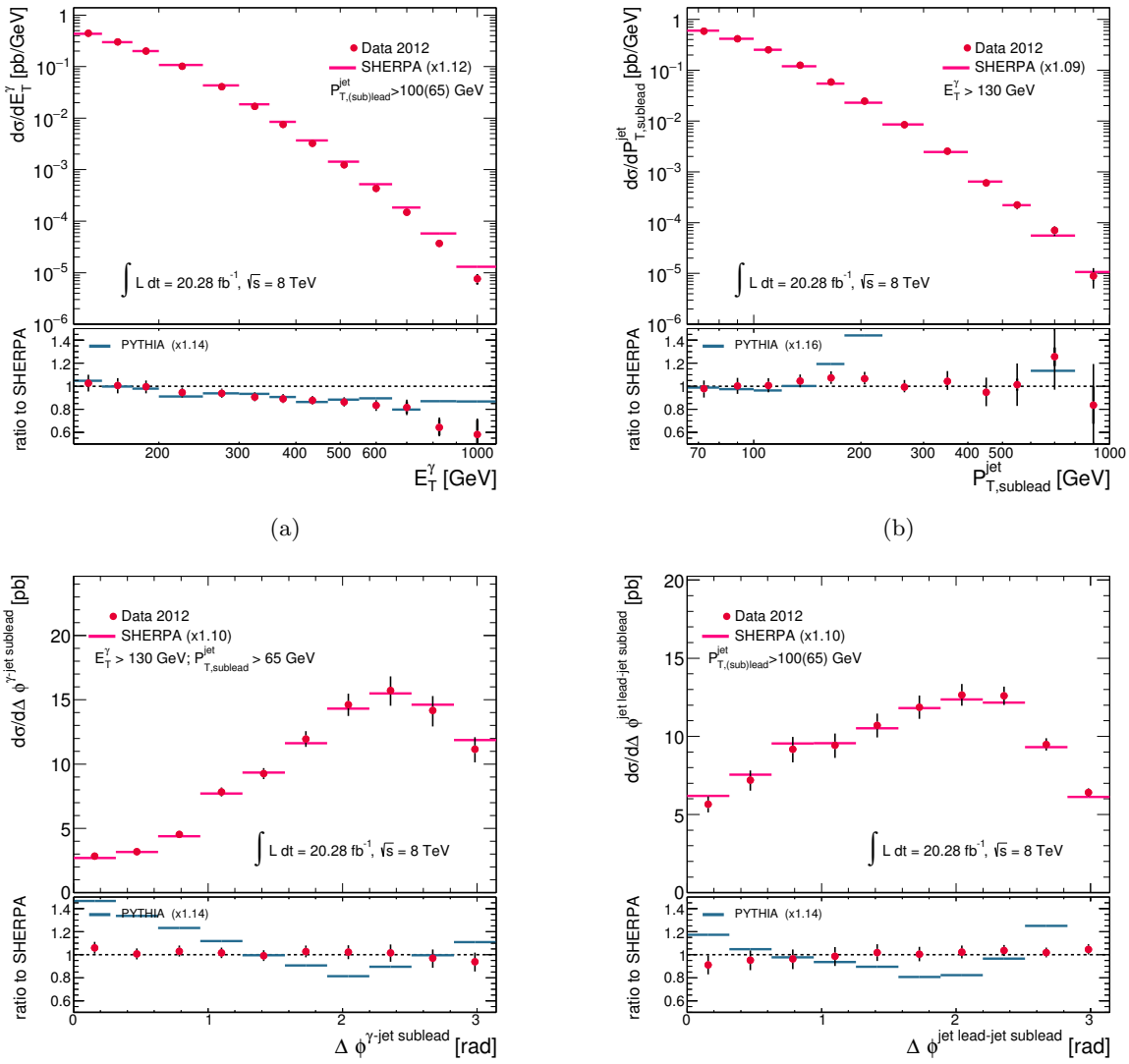
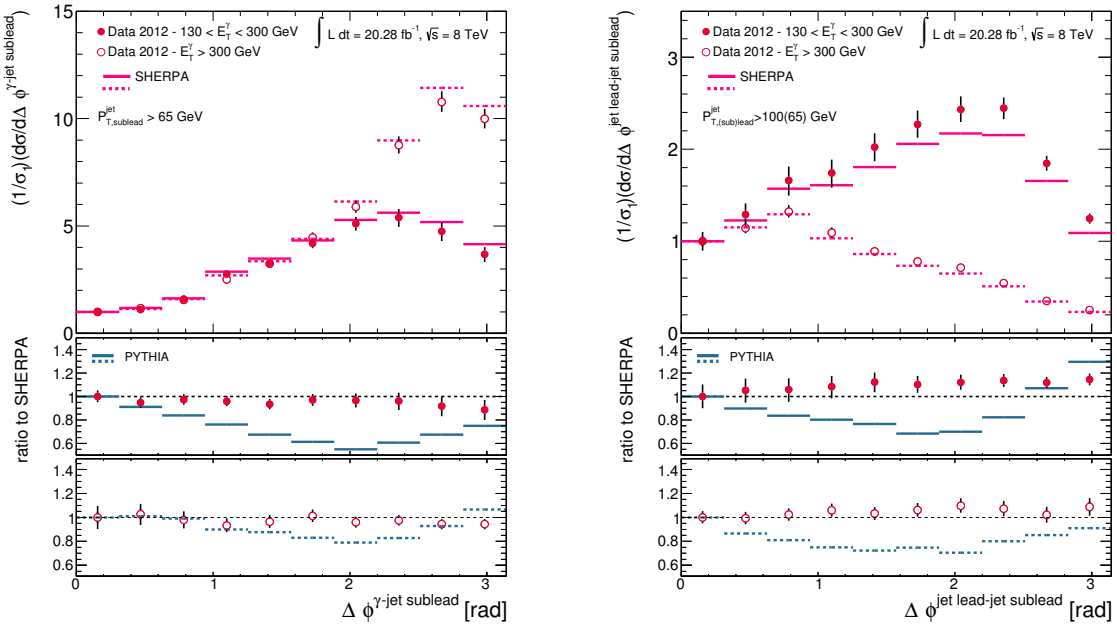


Figure 6.106: *Measured differential cross sections for isolated photon plus two jets production (dots) as functions of* E^{γ}_{T} *(a),* $p^{\mathrm{jet}}_{\mathrm{T,sublead}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$ *(c) and* $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$ *(d). The LO matrix-element plus parton shower predictions from* SHERPA *(pink lines) normalised to the integrated measured cross sections (using the factor indicated) is also shown. The bottom part of each figure shows the ratios of the measured cross section and the prediction of* PYTHIA *(normalised to the measured cross section) to the* SHERPA *prediction. The inner (outer) error bars represent the statistical uncertainties (the statistical and systematic uncertainties added in quadrature). For most of the points, the inner error bars are smaller than the marker size and, thus, not visible.*



(a) (b)

Figure 6.107: *Measured normalised differential cross sections for isolated photon plus two jets production as functions of* $\Delta\phi^{\gamma-\text{jet sublead}}$ *(a) and* $\Delta\phi^{\text{jet lead}-\text{jet sublead}}$ *(b) for* $E_{\text{T}}^{\gamma} < 300$ *GeV (dots) and* $E_{\text{T}}^{\gamma} > 300$ *GeV (open circles). Other details as in the caption to 6.106*

6.13.3 Colour coherence in isolated photon plus two jets production

Figure 6.108a shows the differential cross sections as functions of $\beta^{\text{jet}}_{\text{lead}}$ and β^{γ}. Both cross sections display a very different behaviour: the measured $d\sigma/d\beta^{\text{jet}}_{\text{lead}}$ cross section increases as $\beta^{\text{jet}}_{\text{lead}}$ increases, slower for $\beta^{\text{jet}}_{\text{lead}} < \pi/2$ and faster for $\beta^{\text{jet}}_{\text{lead}} > \pi/2$, whereas $d\sigma/d\beta^{\gamma}$ cross section exhibits a maximum at $\beta^{\gamma} \approx \pi/2$. The predictions of SHERPA provides a good description of the measured cross sections. To ascertain the presence of colour coherence effects from data alone, the ratio of these measured cross sections was performed and is shown in figure 6.108b; enhancements of the cross section as functions of $\beta^{\text{jet}}_{\text{lead}}$ at $\beta^{\text{jet}}_{\text{lead}} = 0$ and at $\beta^{\text{jet}}_{\text{lead}} = \pi$ with respect to the cross section as functions of β^{γ} are clearly observed. This behaviour indicates an enhancement of radiation in the direction of the initial state partons around the leading jet with respect to a non coloured particle, which is consistent with the effect of colour coherence (see section 1.5.4).

Figure 6.109 shows the results as in figure 6.108 but in this case, the range of the leading jet and the photon were restricted to $|\eta^{\text{jet}}_{\text{lead}}| < 1.37$ and $|\eta^{\gamma}| < 1.37$, respectively. The same characteristics as in the full pseudorapidity ranges are observed.

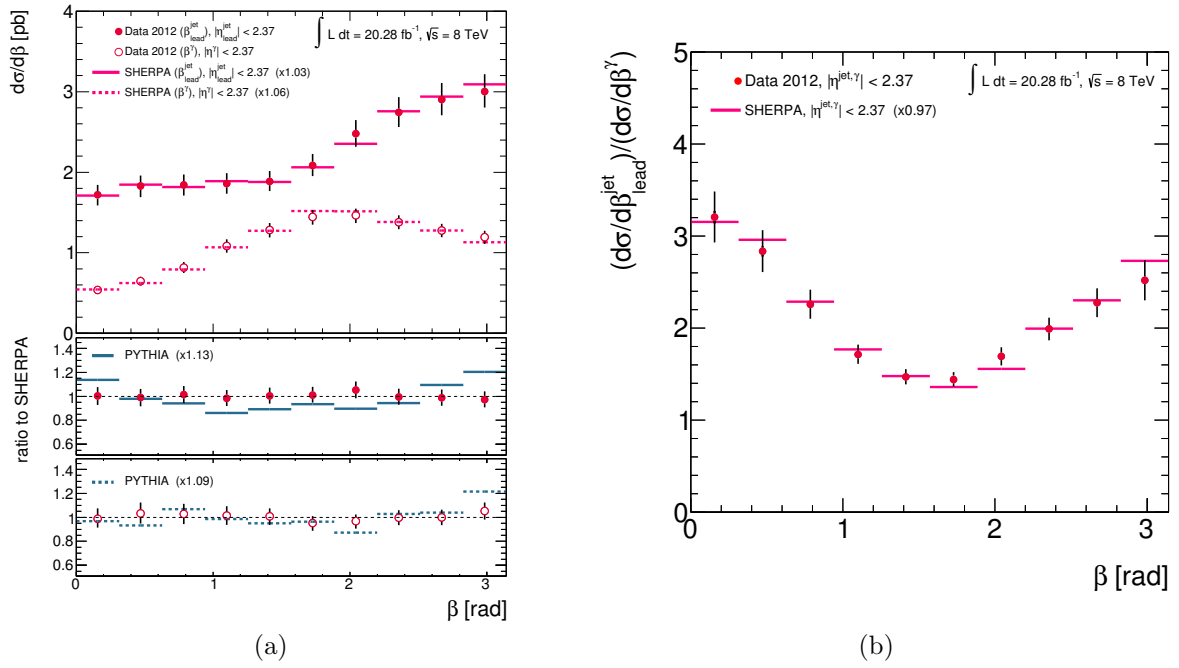
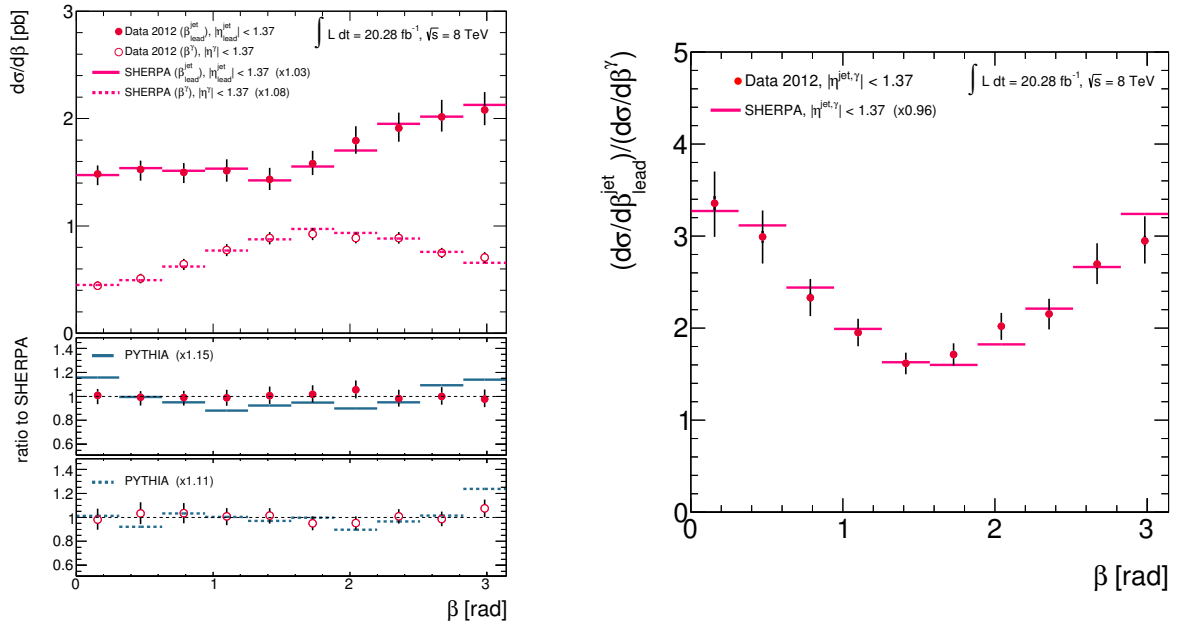


Figure 6.108: *Measured differential cross sections for isolated photon plus two jets production (in the colour coherence samples) as functions of* $\beta^{\text{jet}}_{\text{lead}}$ *(dots) and* β^{γ} *(open circles) for* $|\eta^{jet}_{lead}| < 2.37$ *and* $|\eta^{\gamma}| < 2.37$*, respectively (a). The LO matrix-element plus parton shower predictions from* SHERPA *(pink lines) normalised to the integrated measured cross sections (using the factors indicated in parentheses) are also shown. The bottom part of (a) shows the ratio of the measured cross section and predictions of* PYTHIA *(normalised to the integrated measured cross sections) to the* SHERPA *predictions. In (b) the ratio of the measured* $\beta^{\text{jet}}_{\text{lead}}$ *and* β^{γ} *cross sections (dots) and the ratio predicted by* SHERPA *(pink lines) are shown. Other details as in the caption to 6.106.*



(a) (b)

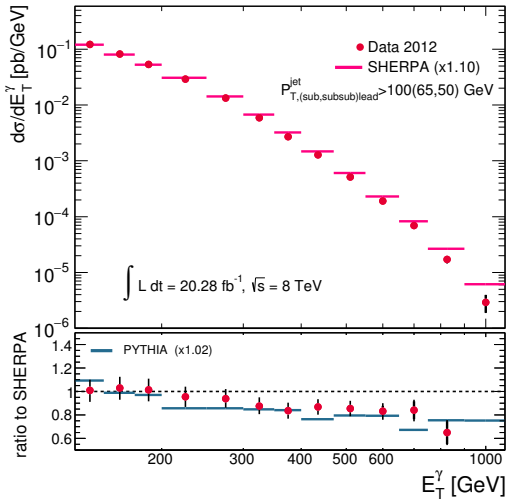
Figure 6.109: *Measured differential cross sections for isolated photon plus two jets production (in the colour coherence samples) as functions of* $\beta^{\text{jet}}_{\text{lead}}$ *(dots) and* β^{γ} *(open circles) for* $|\eta^{jet}_{lead}| < 1.37$ *and* $|\eta^{\gamma}| < 1.37$*, respectively (a). The LO matrix-element plus parton shower predictions from* SHERPA *(pink lines) normalised to the integrated measured cross sections (using the factors indicated in parentheses) are also shown. The bottom part of (a) shows the ratio of the measured cross section and predictions of* PYTHIA *(normalised to the integrated measured cross sections) to the* SHERPA *predictions. In (b) the ratio of the measured* $\beta^{\text{jet}}_{\text{lead}}$ *and* β^{γ} *cross sections (dots) and the ratio predicted by* SHERPA *(pink lines) are shown. Other details as in the caption to 6.106.*

6.13.4 Isolated photon plus three jets production

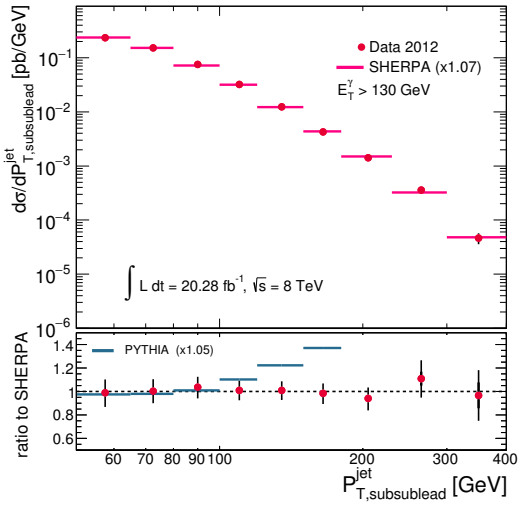
Figure 6.110 shows the isolated photon plus three jets differential cross sections as functions of E^{γ}_{T}, $p^{\text{jet}}_{\text{T,subsublead}}$, $\Delta\phi^{\gamma-\text{jet subsublead}}$, $\Delta\phi^{\text{jet lead}-\text{jet subsublead}}$ and $\Delta\phi^{\text{jet sublead}-\text{jet subsublead}}$. The measured $d\sigma/dE^{\gamma}_{\text{T}}$ and $d\sigma/dp^{\text{jet}}_{\text{T,subsublead}}$ decrease by five and three orders of magnitude, respectively, within the measured range. The LO matrix-element plus parton shower prediction of PYTHIA gives a good description of $d\sigma/dE^{\gamma}_{\text{T}}$, whereas the prediction of SHERPA describes well the measured $d\sigma/dp^{\text{jet}}_{\text{T,subsublead}}$.

The $\Delta\phi^{\gamma-\text{jet subsublead}}$ measured cross sections increases as $\Delta\phi^{\gamma-\text{jet subsublead}}$ increases whereas the measured $\Delta\phi^{\text{jet lead}-\text{jet subsublead}}$ and $\Delta\phi^{\text{jet sublead}-\text{jet subsublead}}$ cross sections are approximately constant. The predictions of SHERPA and PYTHIA describe well the data.

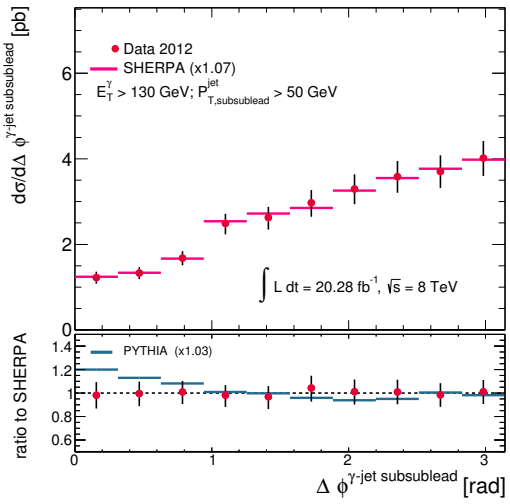
As in the case of the isolated photon plus two jets production, the normalised differential cross section as functions of the azimuthal angles for $E^{\gamma}_{\text{T}} < 300$ GeV and $E^{\gamma}_{\text{T}} > 300$ GeV were measured. Figure 6.111 shows the $\Delta\phi^{\gamma-\text{jet subsublead}}$, $\Delta\phi^{\text{jet lead}-\text{jet subsublead}}$ and $\Delta\phi^{\text{jet sublead}-\text{jet subsublead}}$ normalised differential cross sections for the two E^{γ}_{T} ranges; in this figure, the measured and predicted differential cross sections were normalised to unity at $\Delta\phi = \pi/10$ so that the change in shape at large $\Delta\phi$ values for the two E^{γ}_{T} ranges can be observed. for $E^{\gamma}_{\text{T}} < 300$ GeV the cross sections have similar shape to those measured in the $E^{\gamma}_{\text{T}} > 130$ GeV range; however, for $E^{\gamma}_{\text{T}} > 300$ GeV, the $\Delta\phi^{\gamma-\text{jet subsublead}}$ cross section is enhanced and the $\Delta\phi^{\text{jet lead}-\text{jet subsublead}}$ and $\Delta\phi^{\text{jet sublead}-\text{jet subsublead}}$ cross sections are suppressed for large $\Delta\phi$ values. The predictions of SHERPA describe well the shape of the measured normalised differential cross sections and their evolution with the scale.



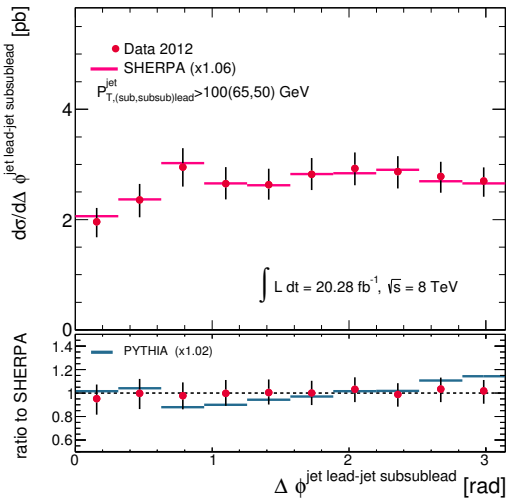
(a)



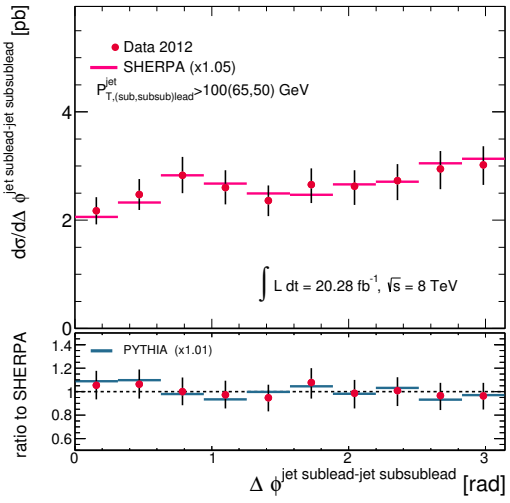
(b)



(c)

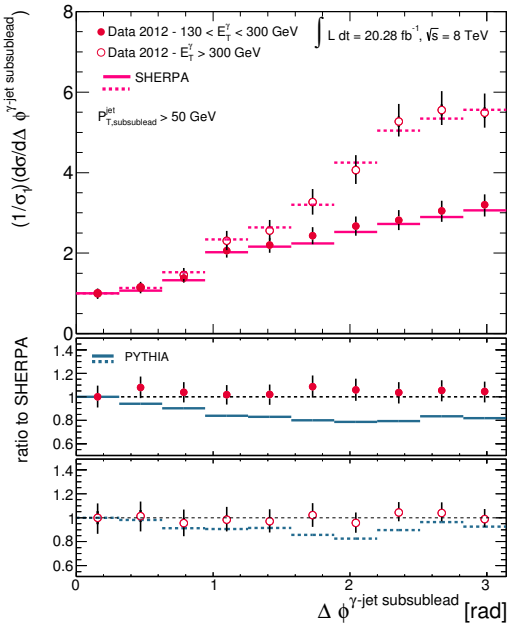


(d)

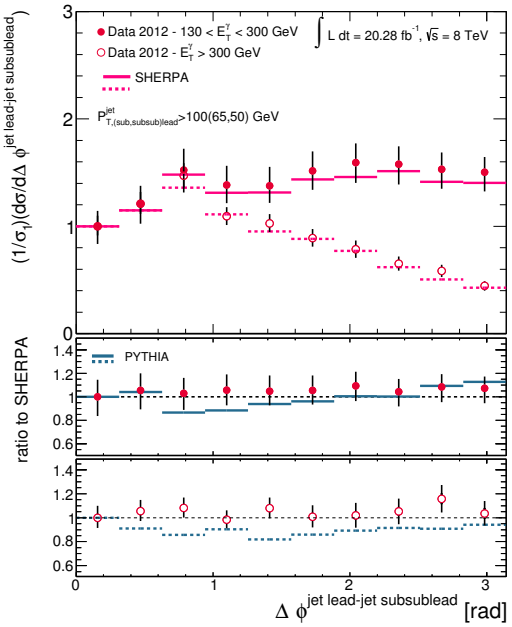


(e)

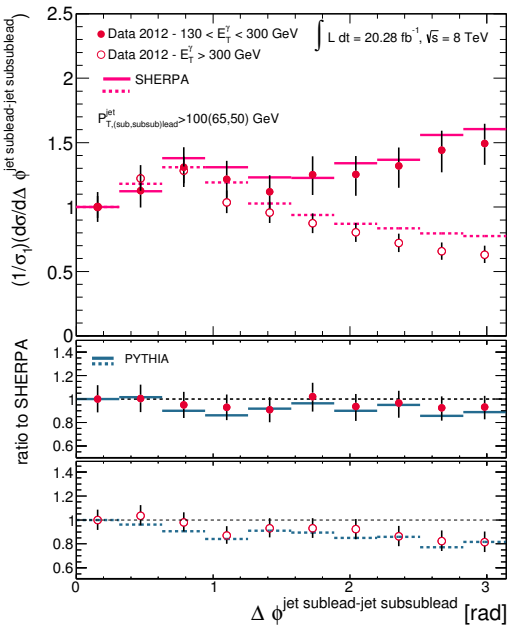
Figure 6.110: *Measured differential cross sections for isolated photon plus three jets production (dots) as functions of* E^{γ}_{T} *(a),* $p^{\mathrm{jet}}_{\mathrm{T,subsublead}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\ subsublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\ lead}-\mathrm{jet\ subsublead}}$ *(d) and* $\Delta\phi^{\mathrm{jet\ sublead}-\mathrm{jet\ subsublead}}$ *(e). Other details as in the caption to figure 6.106.*



(a)



(b)



(c)

Figure 6.111: *Measured normalised differential cross sections for isolated photon plus three jets production as functions of* $\Delta\phi^{\gamma-\text{jet subsublead}}$ *(a),* $\Delta\phi^{\text{jet lead}-\text{jet subsublead}}$ *(b) and* $\Delta\phi^{\text{jet sublead}-\text{jet subsublead}}$ *for* $E_{\mathrm{T}}^{\gamma} < 300$ *GeV (dots) and* $E_{\mathrm{T}}^{\gamma} > 300$ *GeV (open circles). Other details as in the caption to 6.106.*

Chapter 7

Additional studies on photon plus jets events

In this chapter a series of additional studies performed on photon plus jets events in order to support the photon plus jets production analysis of chapter 6 are presented.

7.1 Photon isolation studies

The isolation cut for the event selection ($E_T^{\text{iso}} < 4.8$ GeV $+4.2 \cdot 10^{-3} \cdot E_T^{\gamma}$) was chosen to be E_T^{γ} dependent to keep the signal to background ratio as high as possible as well as a constant reconstruction efficiency throughout the whole E_T^{γ} range. Figure 7.1 shows the efficiency of the E_{T}^{γ}-dependent cut on photon candidates in the signal MC at detector level. The probability of a photon passing the isolation requirement as a function of E_{T}^{γ} is shown for the E_{T}^{γ}-dependent requirement and several fixed cuts. The improved performance of the E_{T}^{γ}-dependent cut is clearly seen, especially at high E_{T}^{γ}, where all fixed thresholds show a sizeable decrease in the probability.

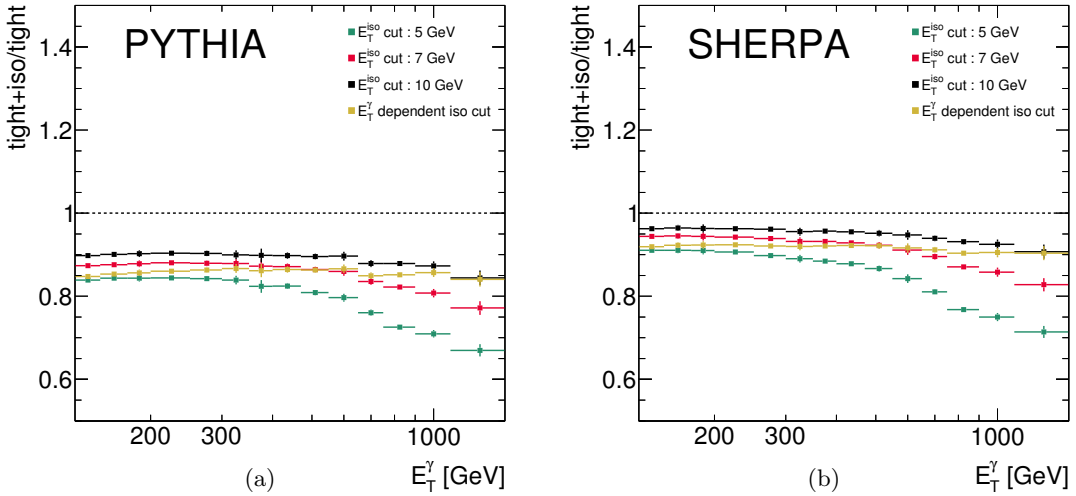


Figure 7.1: *Probability for a photon candidate to pass an isolation requirement with fixed* $E_{\text{T}}^{\text{iso}}$ *values of 5 (green), 7 (red) or 10 (black) GeV. The result for the* E_{T}^{γ}*-dependent requirement is shown as gold squares. The plot in (a) is for* Pythia *MC and the plot in (b) is for* Sherpa *MC.*

The equivalent requirement at particle level was determined by studying the correlation between the detector and particle levels as a function of $E_{\text{T}}^{\text{iso}}$ using photon+one-jet events.

The $E_{\mathrm{T}}^{\mathrm{iso}}$ variable at particle level is computed using stable particles and corrected using the jet-area method (see section 4.4.3). Figure 7.2 shows the (iso+tight)/tight and (iso+part)/part efficiencies at detector and particle levels, respectively, for PYTHIA hard, brem and hard+brem and SHERPA MC samples. In all cases, the detector level requirement depends on E_{T}^{γ}, as explained in the previous paragraph. The particle level efficiency was computed using the same E_{T}^{γ}-dependent requirement and for several fixed E_{T}^{γ} requirements. For both PYTHIA and SHERPA, the detector and particle level efficiencies when using the E_{T}^{γ}-dependent requirement are similar: the difference at low (high) E_{T}^{γ} is 8.9% (2.1%) and 0.5% (4.5%) for PYTHIA and SHERPA, respectively, as can be seen in figure 7.2; these values represent the amount of extrapolation, and its uncertainty arising from signal modelling, due to the isolation requirement when unfolding the E_{T}^{γ}-dependent requirement from detector to particle level.

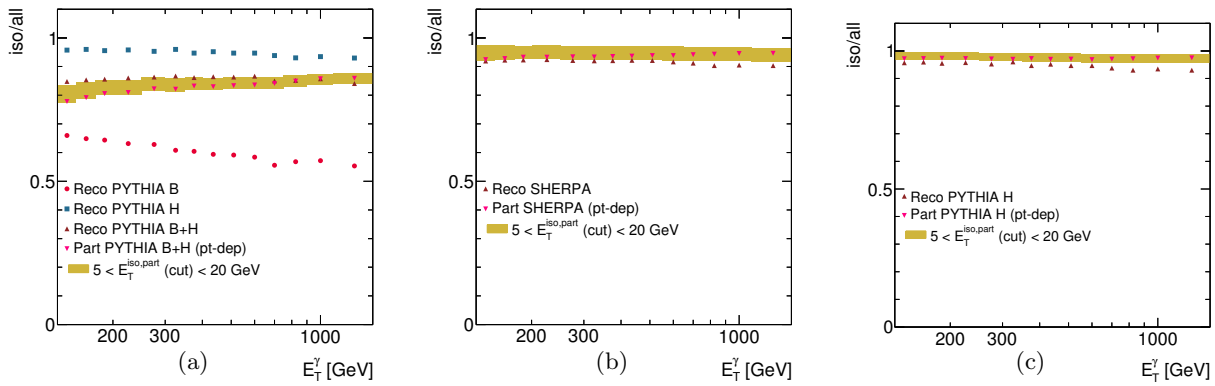
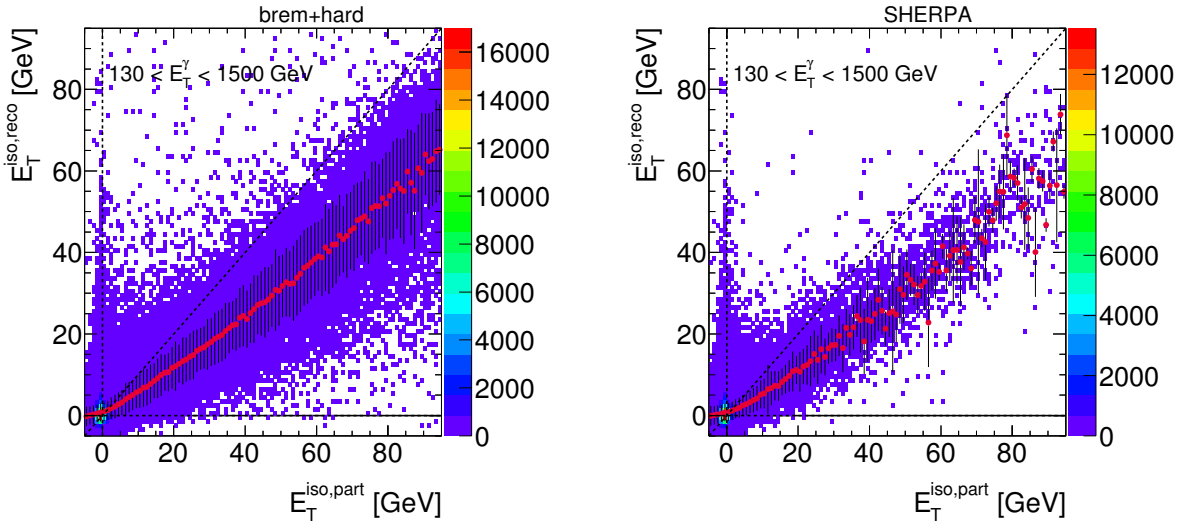


Figure 7.2: *Probability for a photon candidate to pass the isolation requirement at detector level (upwards triangles) as a function of* E_{T}^{γ} *for* PYTHIA *(a) hard+brem and* SHERPA *(b) using the* E_{T}^{γ}*-dependent isolation requirement. The probability at particle level, also using a* E_{T}^{γ}*-dependent requirement is displayed as downwards triangles in these figures. Figure (a) also includes the individual hard (squares) and brem (dots) probabilities at detector level for* PYTHIA*. Figure (c) shows the probability for* PYTHIA *hard only both at detector and particle levels. In these three figures, the shaded band represents the probability for a fixed cut in* E_{T}^{γ} *varying from 5 to 20 GeV at particle level.*

To determine which requirement at particle level would provide the smallest extrapolation, studies of the correlation between isolation energy at detector and particle levels were performed and explained below.

Figure 7.3 shows the correlation between the detector and particle levels for the $E_{\mathrm{T}}^{\mathrm{iso}}$ variable. In both PYTHIA and SHERPA there is a very good correlation in the signal region; in the tail of the distribution (which is dominated by background in the data) the correlation is situated below the diagonal and presents a wide spread. The profiles of the correlations are also included in figure 7.3, which are used in the following discussion to make comparisons.

Figure 7.4 shows the profiles for PYTHIA and SHERPA for $130 < E_{\mathrm{T}}^{\gamma} < 150$ GeV and $750 < E_{\mathrm{T}}^{\gamma} < 900$ GeV. These plots confirm that there is a dependence at the detector level with E_{T}^{γ} in the signal region. The same trend is observed in the signal region for both MC samples and for PYTHIA hard and brem, separately, as can be seen in the comparisons of figures 7.5 and 7.6. Thus, the same behaviour is seen for PYTHIA hard, PYTHIA brem and SHERPA in each E_{T}^{γ} region.



(a) (b)

Figure 7.3: *The* $(E_{\mathrm{T}}^{\mathrm{iso}})^{\mathrm{reco}}$ *vs.* $(E_{\mathrm{T}}^{\mathrm{iso}})^{\mathrm{part}}$ *correlation (blue dots) for* PYTHIA *(a) hard+brem and* SHERPA *(b) for* $130 < E_{\mathrm{T}}^{\gamma} < 1500$ *GeV. The profile of the correlation is displayed as red dots; the error bars represent the width of the distribution in each bin.*

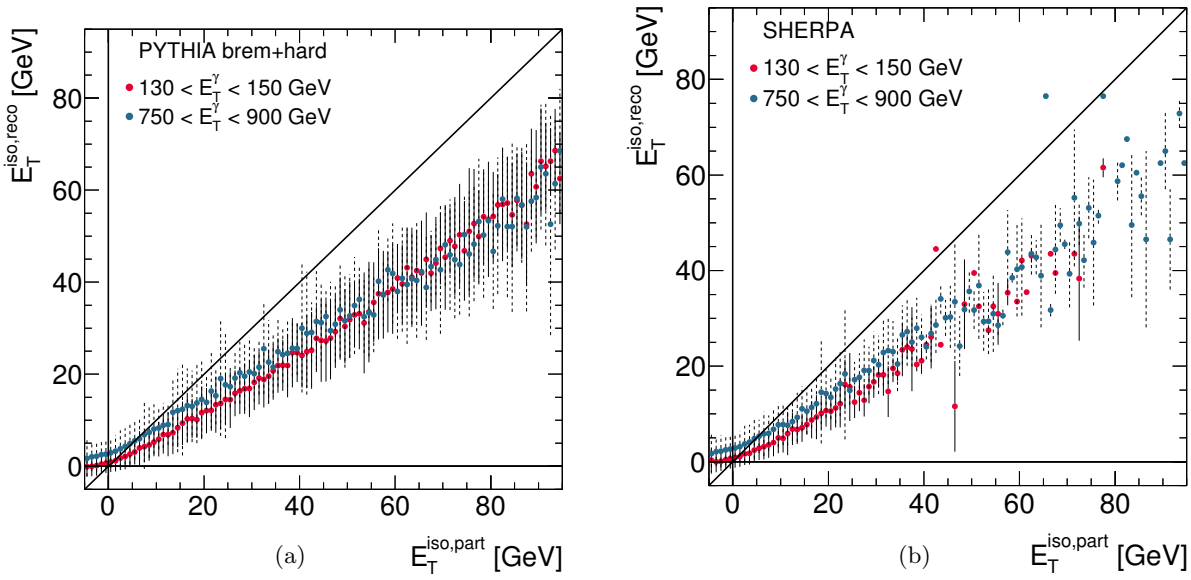


Figure 7.4: *The profile of the* $(E_{\mathrm{T}}^{\mathrm{iso}})^{\mathrm{reco}}$ *vs.* $(E_{\mathrm{T}}^{\mathrm{iso}})^{\mathrm{part}}$ *correlation (dots) for* PYTHIA *hard+brem (a) and* SHERPA *(b) in different* E_{T}^{γ} *regions.*

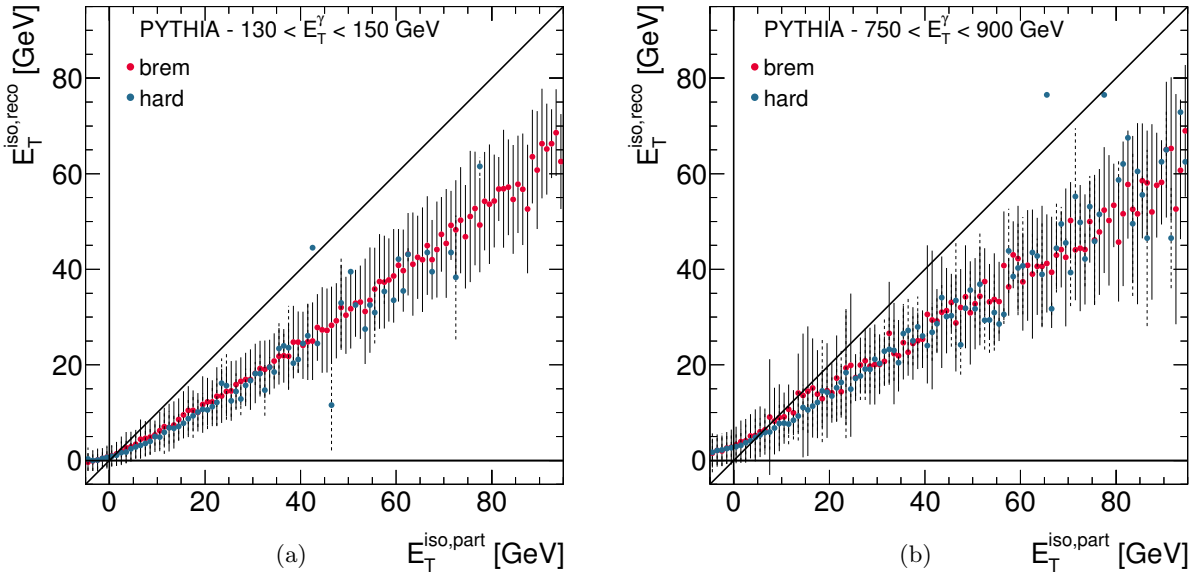


Figure 7.5: *The profile of the* $(E_{\mathrm{T}}^{\mathrm{iso}})^{\mathrm{reco}}$ *vs.* $(E_{\mathrm{T}}^{\mathrm{iso}})^{\mathrm{part}}$ *correlation (dots) for* PYTHIA *hard and brem separately for* $130 < E_{\mathrm{T}}^{\gamma} < 150$ *GeV (a) and* $750 < E_{\mathrm{T}}^{\gamma} < 900$ *GeV (b).*

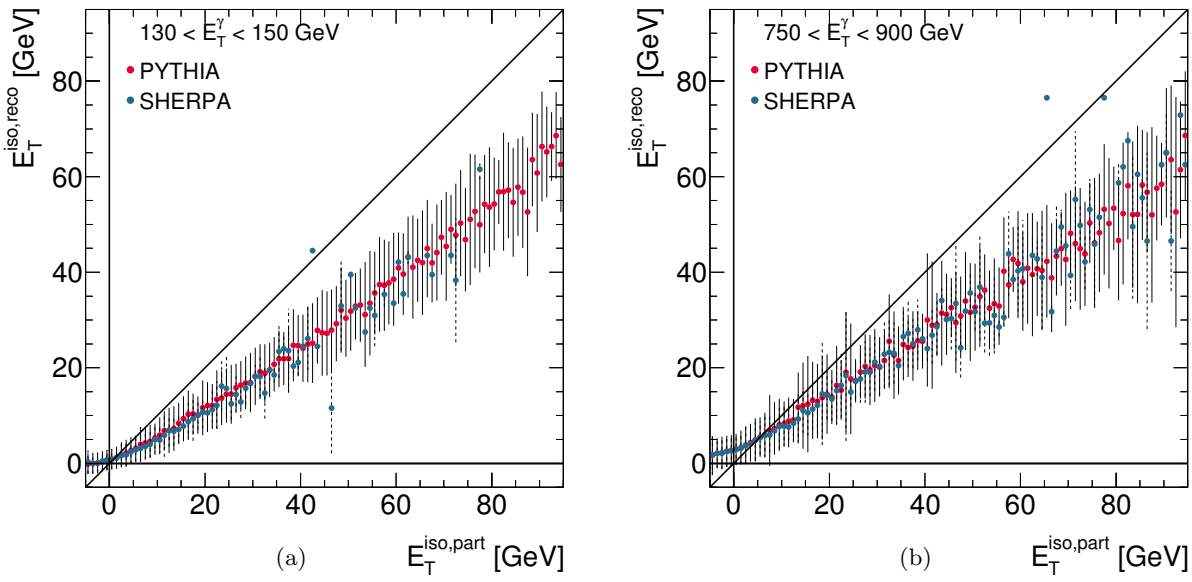


Figure 7.6: *The profile of the* $(E_{\mathrm{T}}^{\mathrm{iso}})^{\mathrm{reco}}$ *vs.* $(E_{\mathrm{T}}^{\mathrm{iso}})^{\mathrm{part}}$ *correlation (dots) for* PYTHIA *hard+brem and* SHERPA *for* $130 < E_{\mathrm{T}}^{\gamma} < 150$ *GeV (a) and* $750 < E_{\mathrm{T}}^{\gamma} < 900$ *GeV (b).*

The requirement on $E_{\mathrm{T}}^{\mathrm{iso}}$ at particle level equivalent to the E_{T}^{γ}-dependent one applied at detector level was determined by performing a χ^2 fit to the profiles in the region $5 < E_{\mathrm{T}}^{\mathrm{iso}} < 20$ GeV in each E_{T}^{γ} region. The fits are displayed in figures 7.7 and 7.8 for PYTHIA and SHERPA, respectively, in each E_{T}^{γ} region. A summary of the values obtained for the requirement at particle level as a function of E_{T}^{γ} is shown in figure 7.9. A χ^2 fit to the results obtained gives 9.59 ± 0.07 (9.68 ± 0.07) GeV for PYTHIA (SHERPA). Therefore, a value of 10 GeV was used for the $E_{\mathrm{T}}^{\mathrm{iso}}$ requirement at particle level.

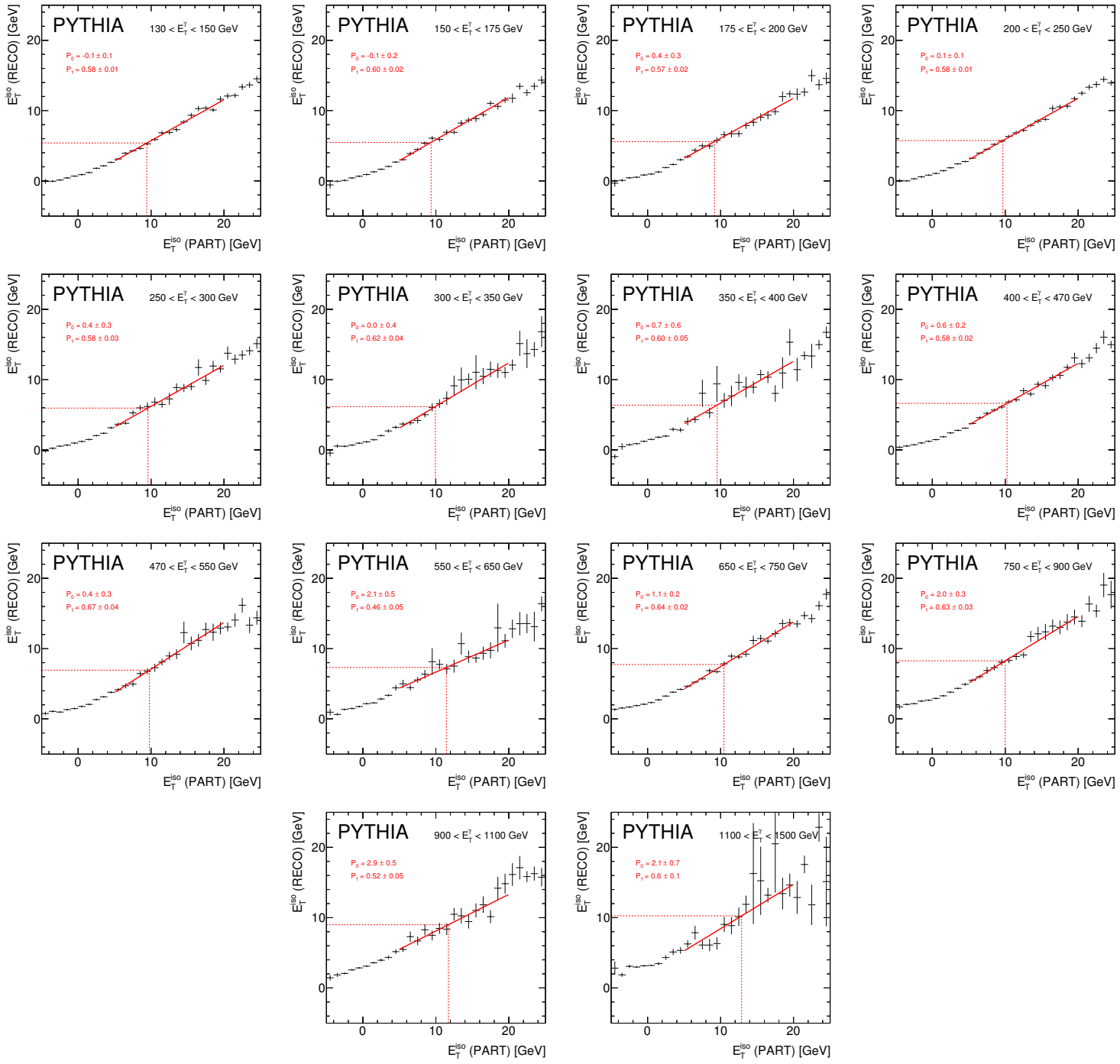


Figure 7.7: $E_{\mathrm{T}}^{\mathrm{iso}}$ *detector vs. particle profiles in different* E_{T}^{γ} *regions (dots) for* PYTHIA *hard+brem. The solid line displays the* χ^2 *fit performed to the profile. The horizontal dashed line corresponds to the* $E_{\mathrm{T}}^{\mathrm{iso}}$ *requirement at detector level, whereas the vertical dashed line is the extrapolated requirement at particle level.*

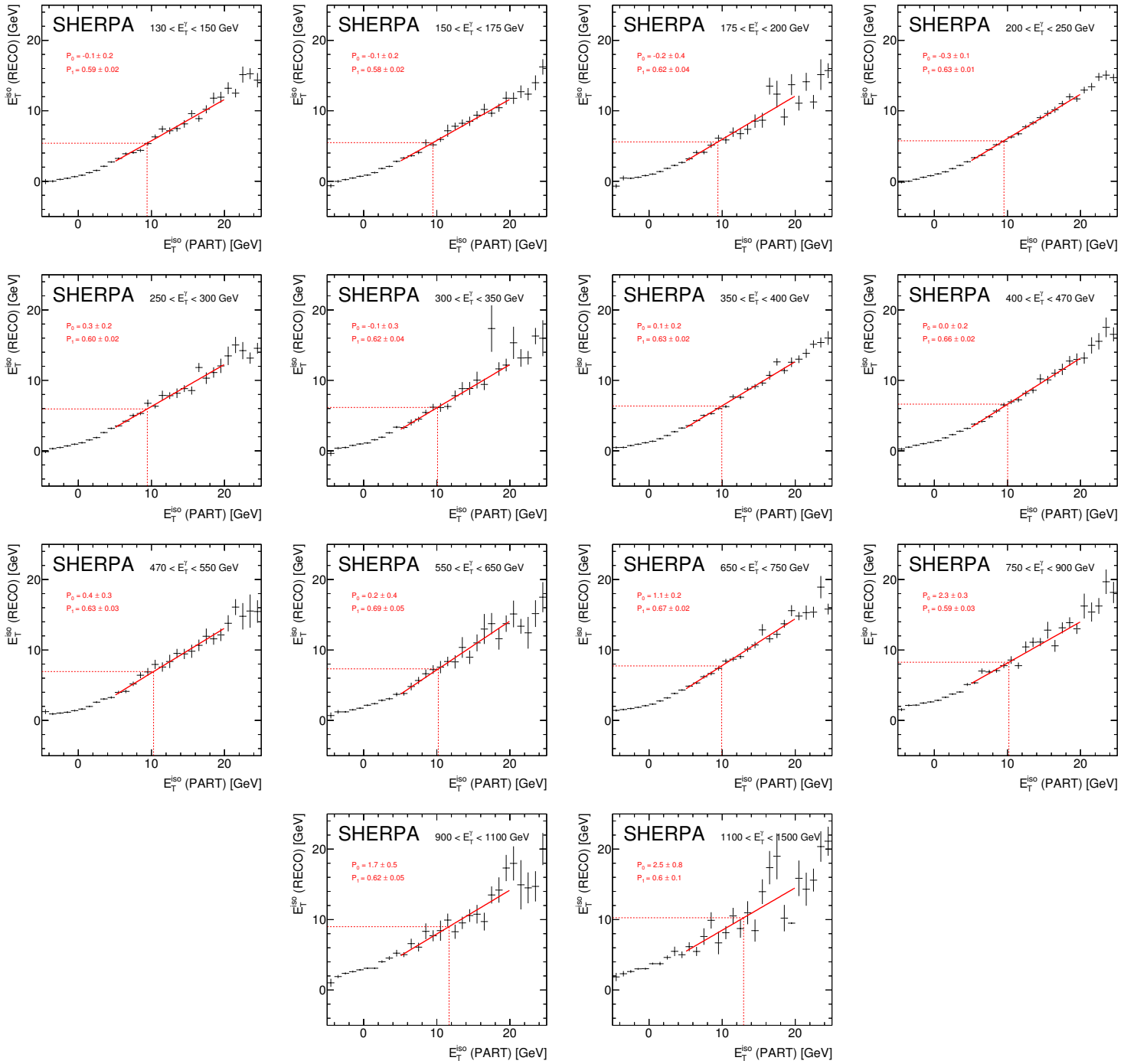


Figure 7.8: $E_{\mathrm{T}}^{\mathrm{iso}}$ *detector vs. particle profiles in different* E_{T}^{γ} *regions (dots) for* SHERPA. *The solid line displays the* χ^2 *fit performed to the profile. The horizontal dashed line corresponds to the* $E_{\mathrm{T}}^{\mathrm{iso}}$ *requirement at detector level, whereas the vertical dashed line is the extrapolated requirement at particle level.*

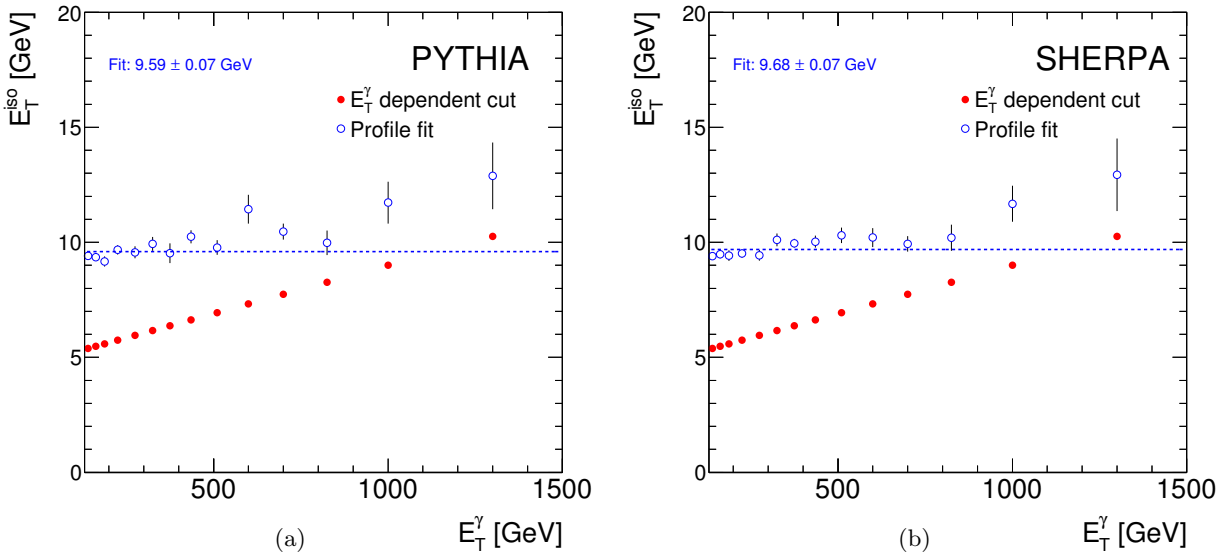


Figure 7.9: *Summary of the values obtained for the requirement at particle level as a function of* E_{T}^{γ} *for* PYTHIA *(a) and* SHERPA *(b) (open circles). The dots represent an* E_{T}^{γ}*-dependent requirement in the center of each bin. The dashed line is the result of a straight-line fit to the the values obtained for the requirement at particle level.*

7.2 Trigger efficiency study

The efficiency of the trigger chain used in the photon+jets analysis (EF_g120_loose) was evaluated from data using a lower threshold prescaled chain. Specifically, the efficiency was evaluated using the EF_g100_loose trigger chain as:

$$\epsilon_{\text{trigger}} = \frac{N_{\text{EF\_g100\_loose}\otimes\text{EF\_g120\_loose}}}{N_{\text{EF\_g100\_loose}}}, \tag{7.1}$$

where $N_{\text{EF\_g100\_loose}\otimes\text{EF\_g120\_loose}}$ is the number of events that are triggered by both chains and $N_{\text{EF\_g100\_loose}}$ is the number of events that are triggered by the EF_g100_loose chain. Only good quality events passing the primary vertex criteria (see section 6.3) were used.

The resulting efficiencies of the EF_g120_loose chain for $E_{\text{T}}^{\gamma} > 130$ GeV photons are:

- *loose prime* selection: 99.84%
- *loose prime* and isolated selection: 99.85%
- *tight* selection: 99.87%
- *tight* and isolated selection (signal region): 99.87%

Figure 7.10 shows the trigger efficiency as a function of E_{T}^{γ} for the four selections listed above.

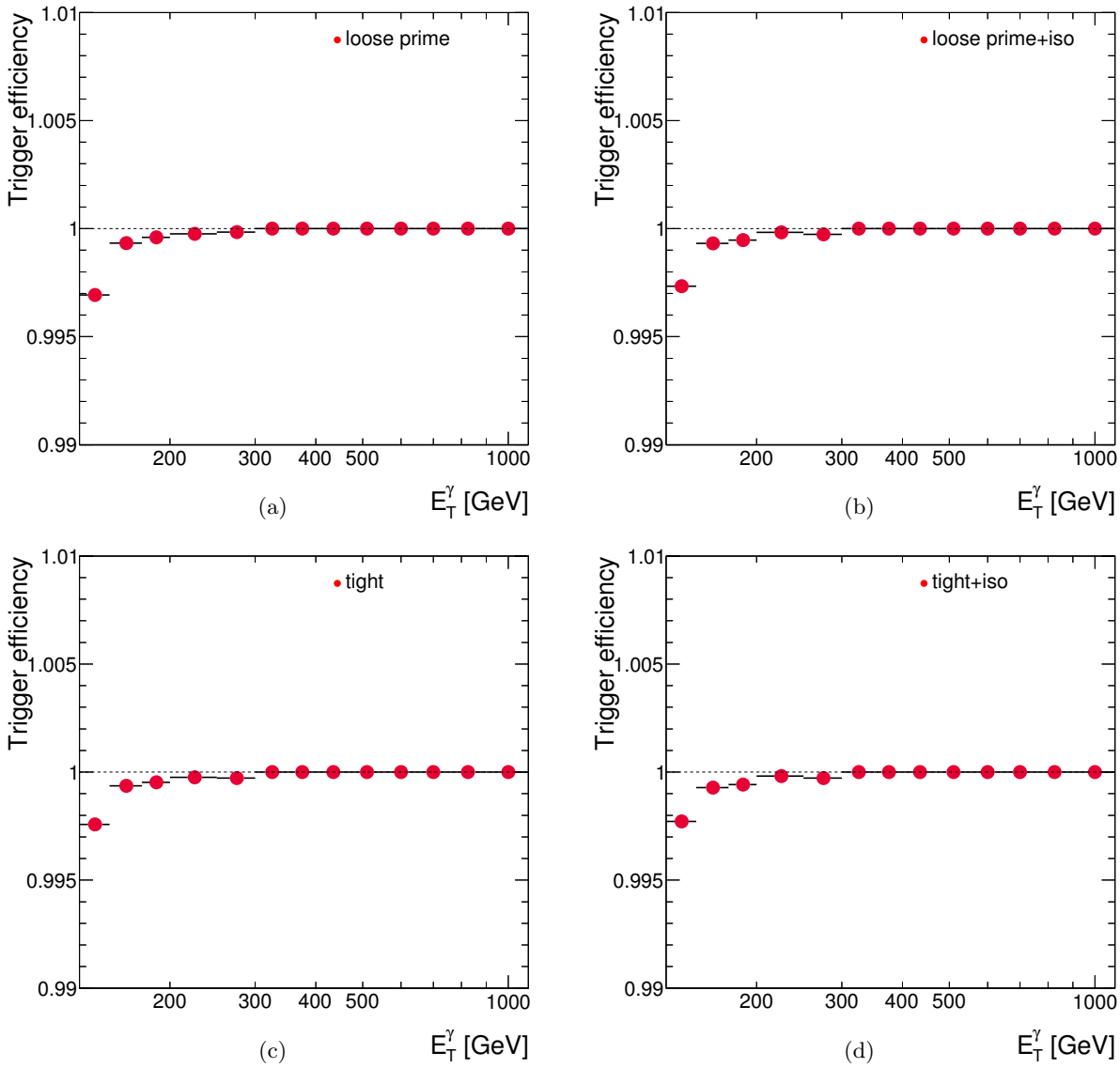
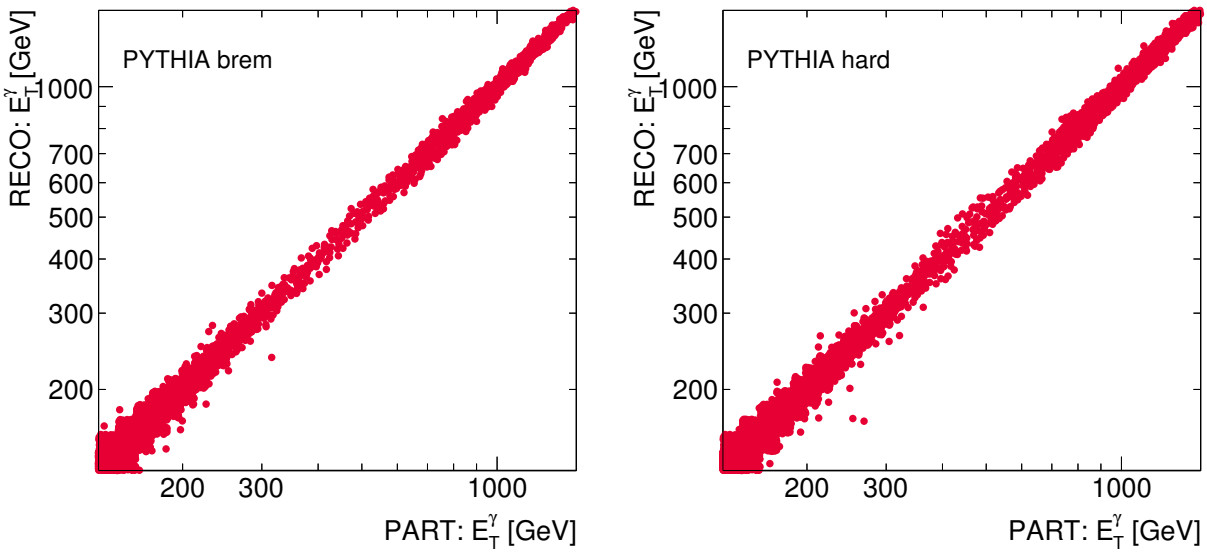


Figure 7.10: *Trigger efficiency as a function of E_T^γ for loose prime (a), loose prime isolated (b), tight (c) and tight isolated (d) selections.*

7.3 Reconstruction quality

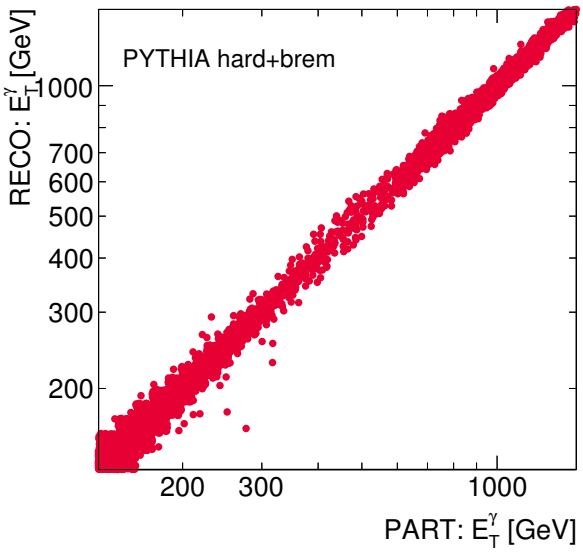
The signal quality reconstruction in photon+jets events was evaluated using the MC samples. The observables at detector and particle levels were compared in an event-by-event basis. A MC event was required to fullfill both the selection criteria at detector and particle levels (see section 6.3); the detector and particle level jets were required to be matched using $\Delta R = 0.6$.

Figures 7.11 to 7.25 show the correlation between the detector and particle levels for the observables studied using the samples of PYTHIA and SHERPA. A very good reconstruction quality is obtained in all cases.

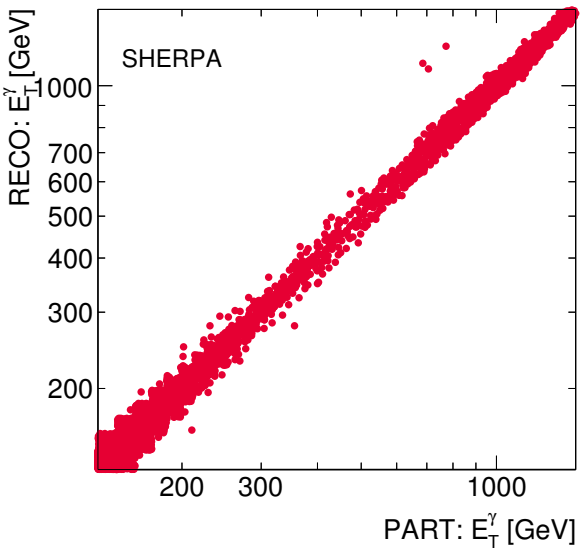


(a)

(b)

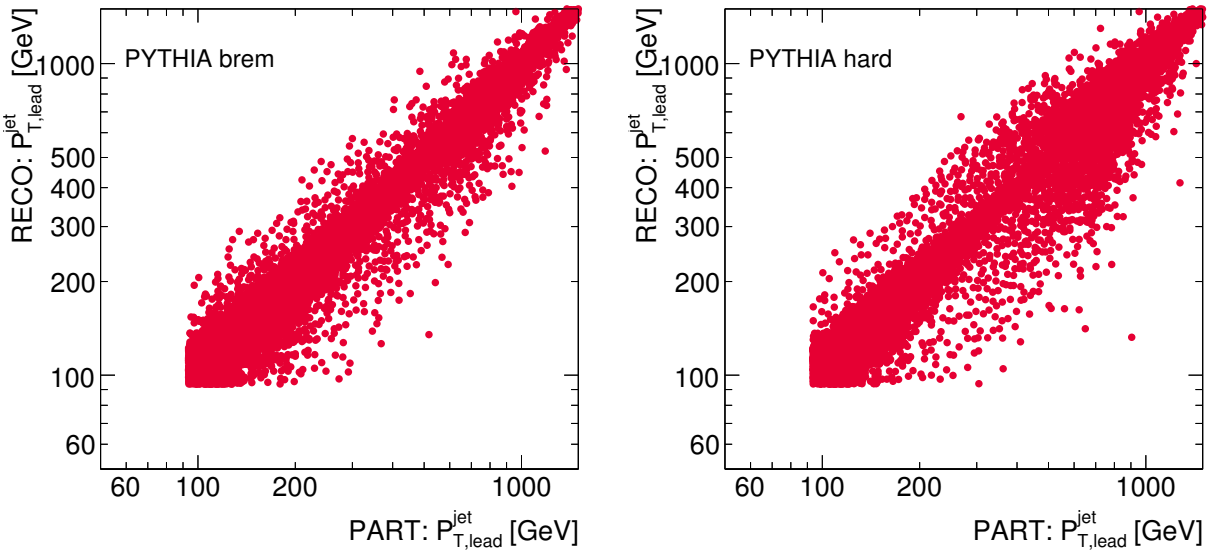


(c)



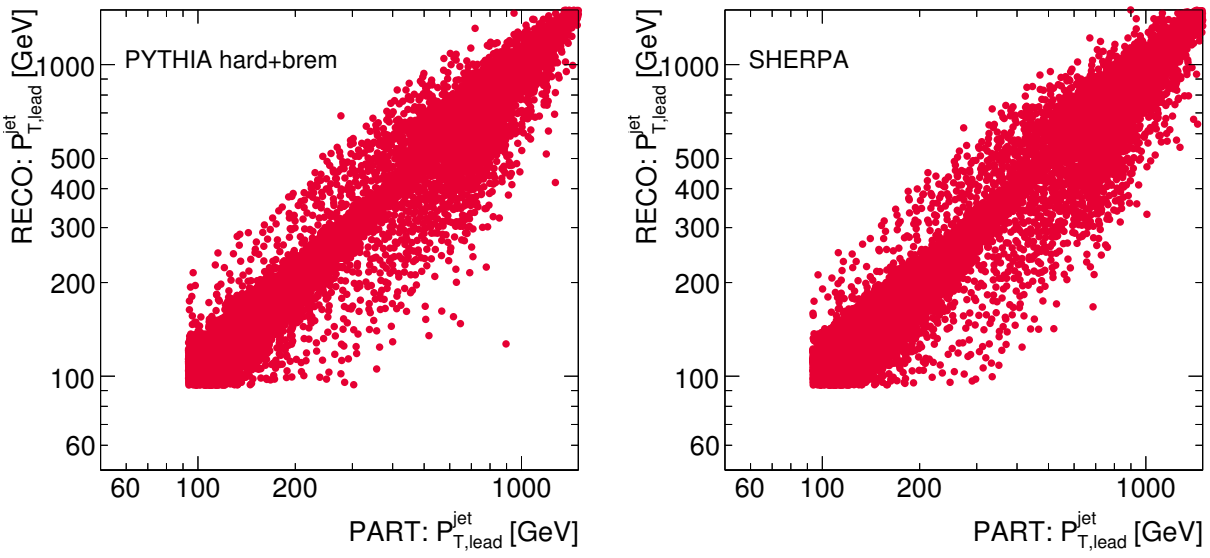
(d)

Figure 7.11: *The* $(E_{\mathrm{T}}^{\gamma})_{\mathrm{reco}}$ *vs.* $(E_{\mathrm{T}}^{\gamma})_{\mathrm{part}}$ *correlation for the photon+one-jet sample for* PYTHIA *brem (a),* PYTHIA *hard (b),* PYTHIA *(c) brem+hard and* SHERPA *(d).*



(a)

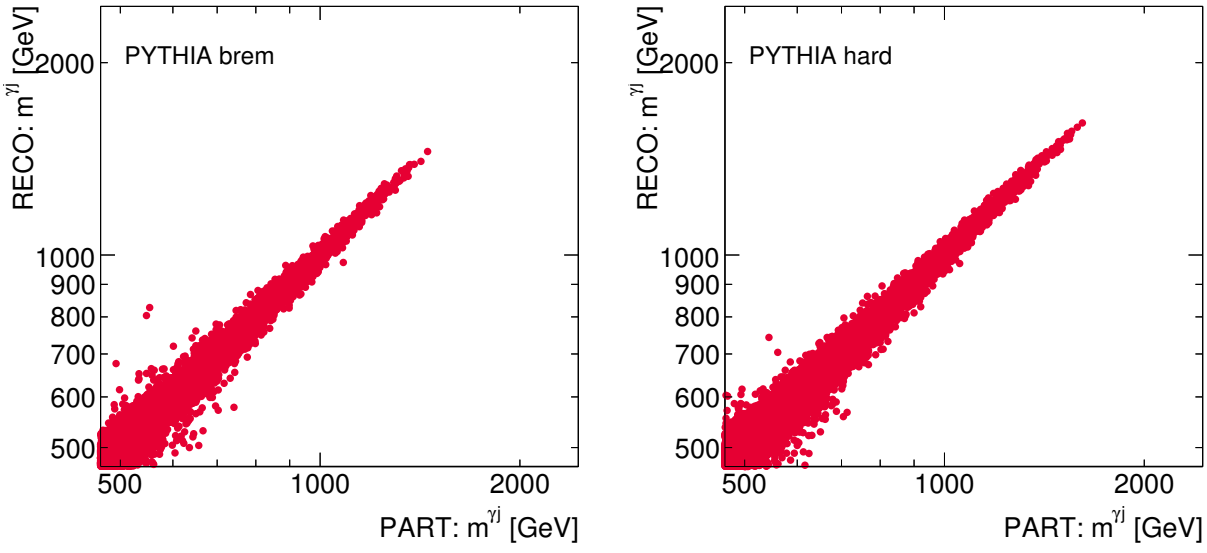
(b)



(c)

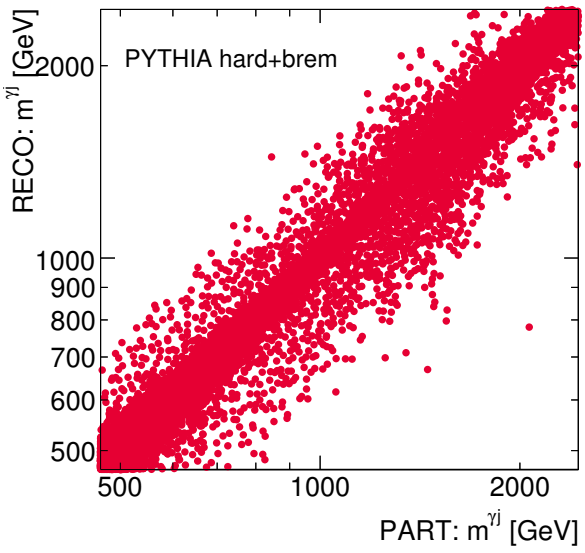
(d)

Figure 7.12: *The* $(p_{\mathrm{T,lead}}^{\mathrm{jet}})_{\mathrm{reco}}$ *vs.* $(p_{\mathrm{T,lead}}^{\mathrm{jet}})_{\mathrm{part}}$ *correlation for the photon+one-jet sample for* PYTHIA *brem (a),* PYTHIA *hard (b),* PYTHIA *(c) brem+hard and* SHERPA *(d).*

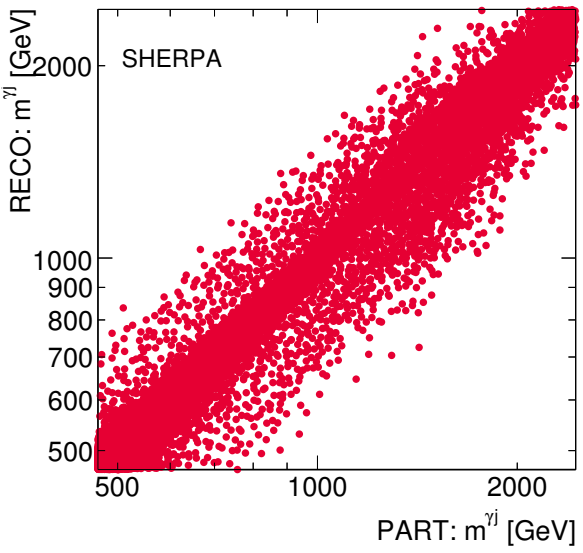


(a)

(b)

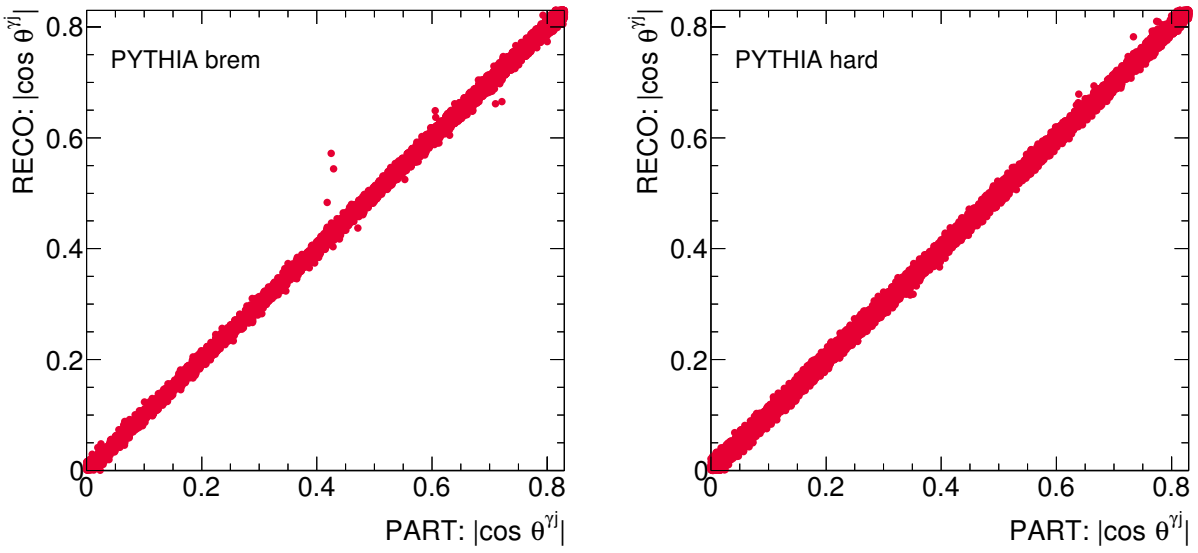


(c)



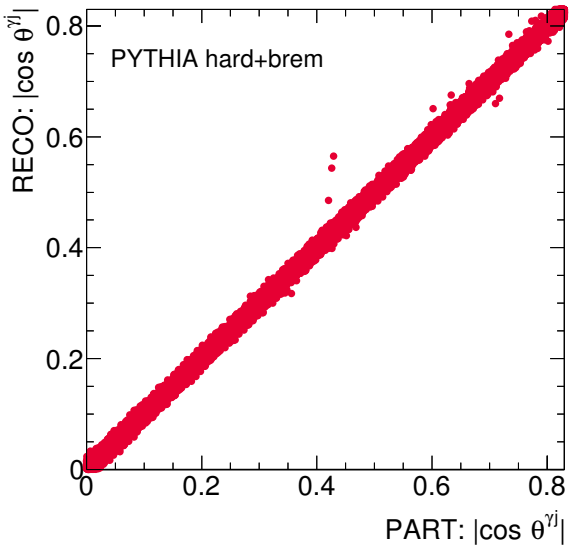
(d)

Figure 7.13: *The* $(m^{\gamma j})_{\text{reco}}$ *vs.* $(m^{\gamma j})_{\text{part}}$ *correlation for the photon+one-jet sample for* PYTHIA *brem (a),* PYTHIA *hard (b),* PYTHIA *(c) brem+hard and* SHERPA *(d).*

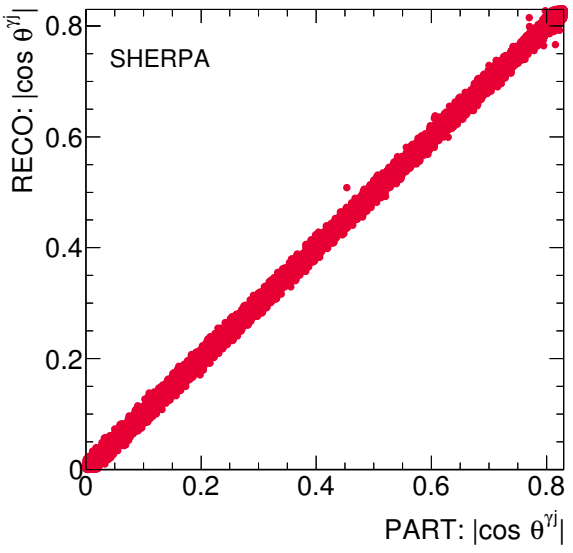


(a)

(b)

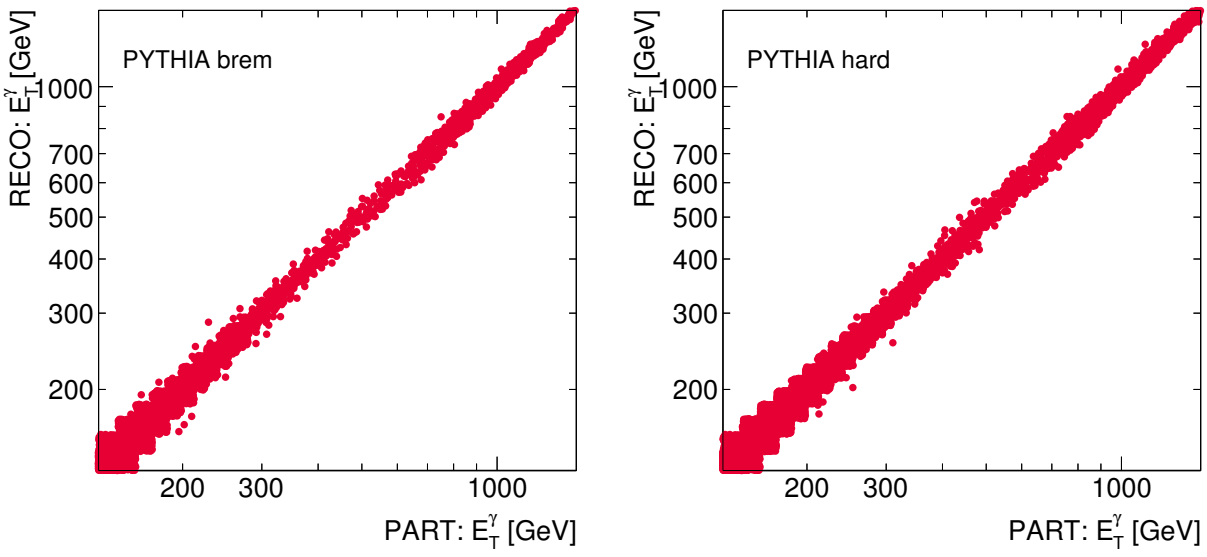


(c)



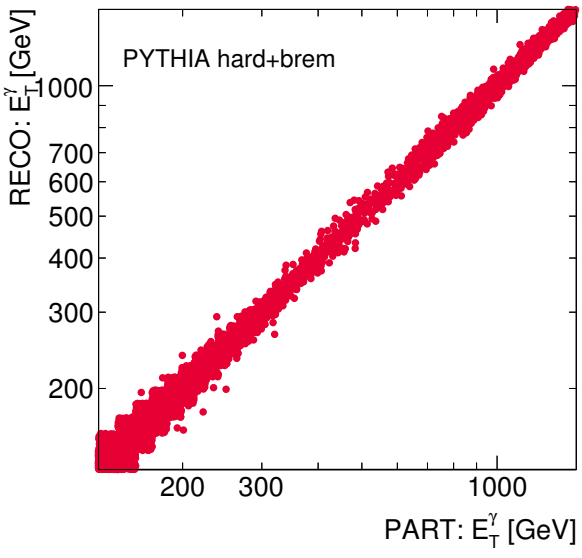
(d)

Figure 7.14: *The* $(\cos\theta^{\gamma j})_{\mathrm{reco}}$ *vs.* $(\cos\theta^{\gamma j})_{\mathrm{part}}$ *correlation for the photon+one-jet sample for* PYTHIA *brem (a),* PYTHIA *hard (b),* PYTHIA *(c) brem+hard and* SHERPA *(d).*

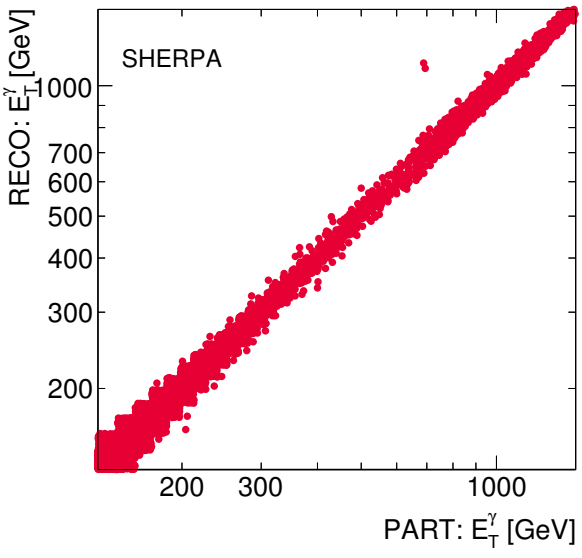


(a)

(b)

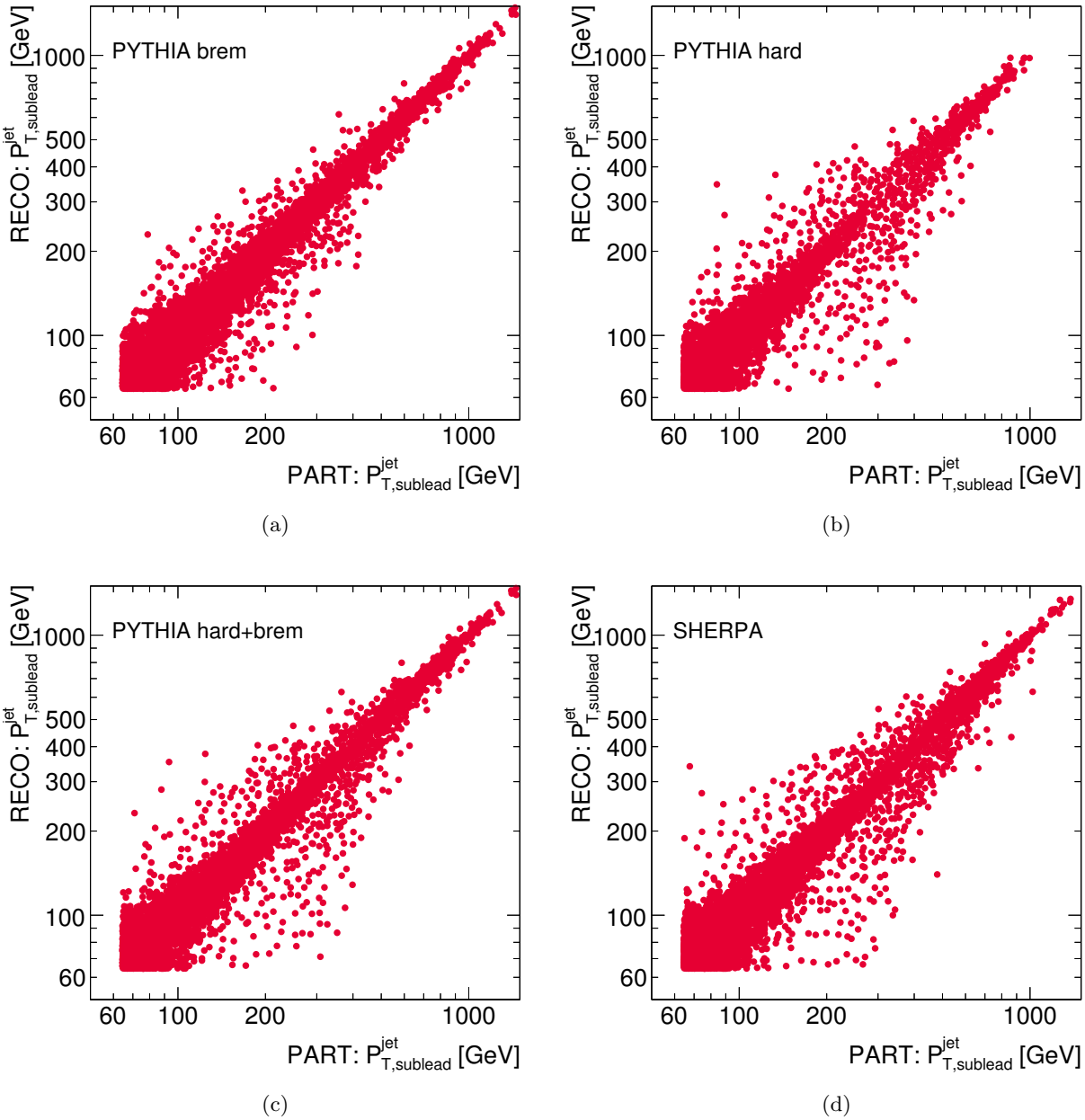


(c)



(d)

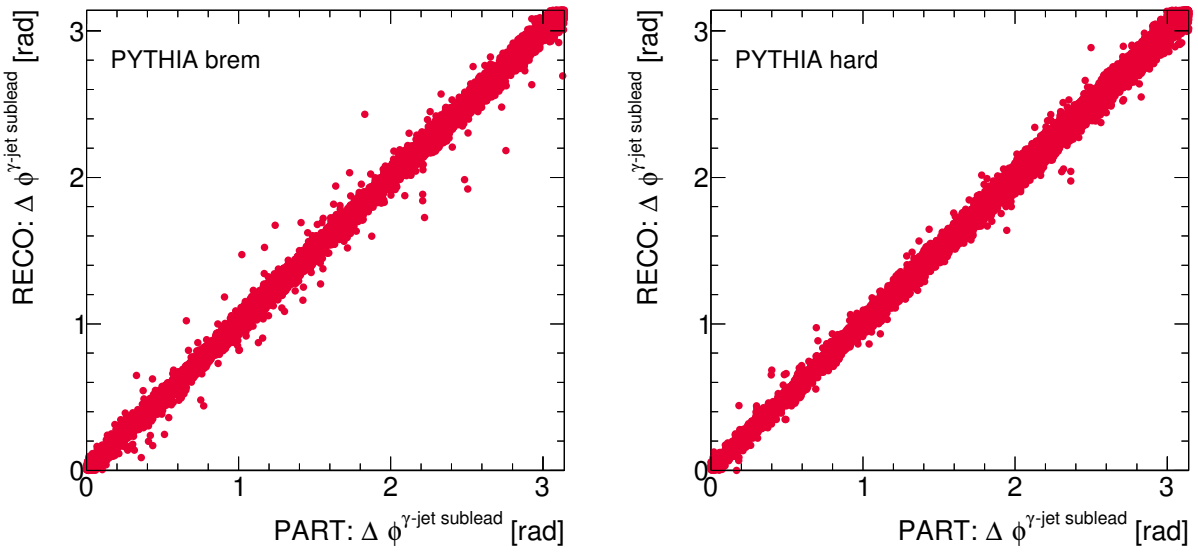
Figure 7.15: *The* $(E_{\mathrm{T}}^{\gamma})_{\mathrm{reco}}$ *vs.* $(E_{\mathrm{T}}^{\gamma})_{\mathrm{part}}$ *correlation for the photon+two-jets sample for* PYTHIA *brem (a),* PYTHIA *hard (b),* PYTHIA *(c) brem+hard and* SHERPA *(d).*



(a) (b)

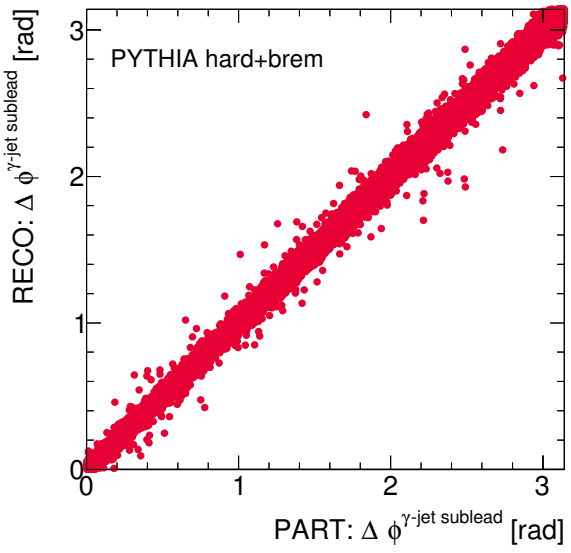
(c) (d)

Figure 7.16: *The* $(p^{\text{jet}}_{\text{T,sublead}})_{\text{reco}}$ *vs.* $(p^{\text{jet}}_{\text{T,sublead}})_{\text{part}}$ *correlation for the photon+two-jets sample for* PYTHIA *brem (a),* PYTHIA *hard (b),* PYTHIA *(c) brem+hard and* SHERPA *(d).*

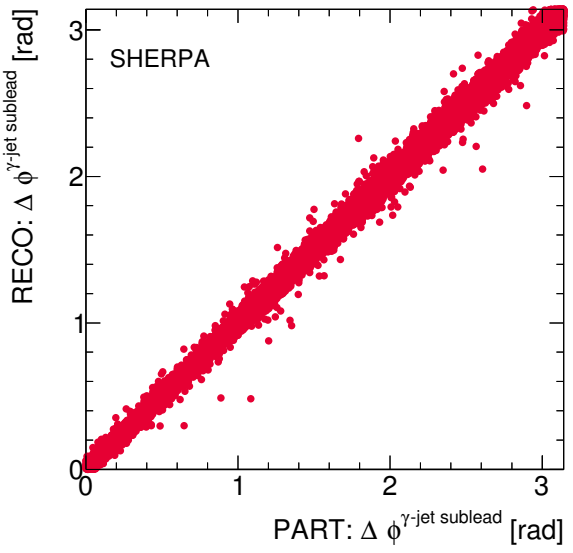


(a)

(b)

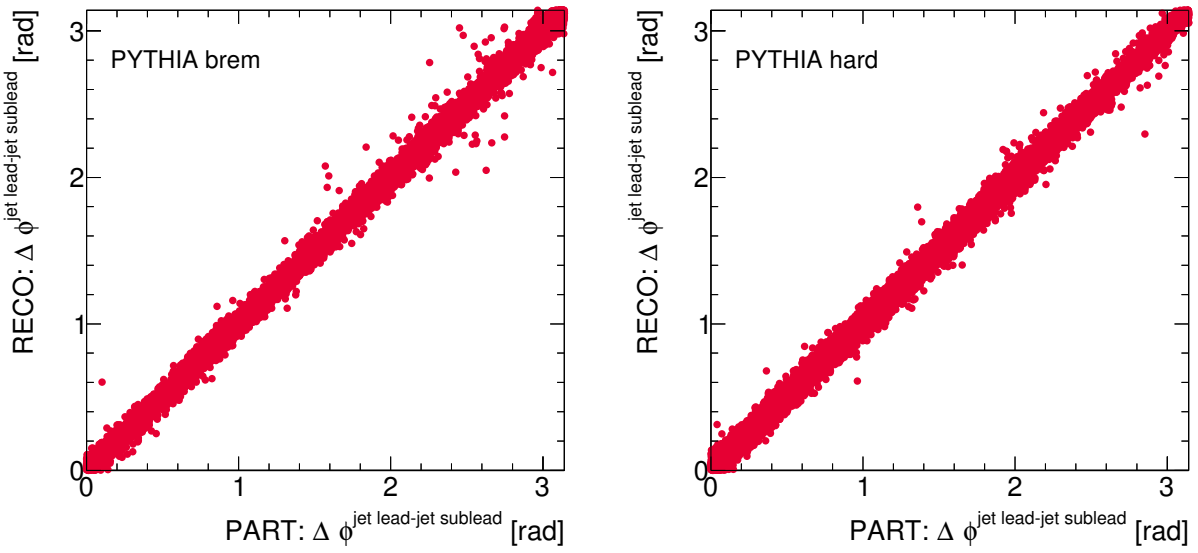


(c)



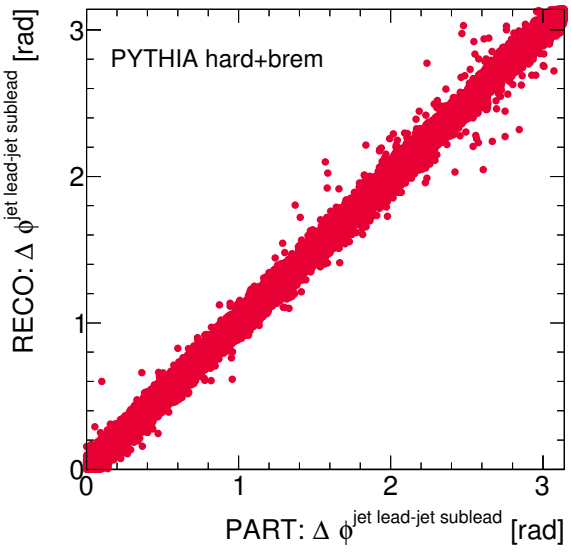
(d)

Figure 7.17: *The* $(\Delta\phi^{\gamma-\text{jet sublead}})_{\text{reco}}$ *vs.* $(\Delta\phi^{\gamma-\text{jet sublead}})_{\text{part}}$ *correlation for the photon+two-jets sample for* PYTHIA *brem (a),* PYTHIA *hard (b),* PYTHIA *(c) brem+hard and* SHERPA *(d).*

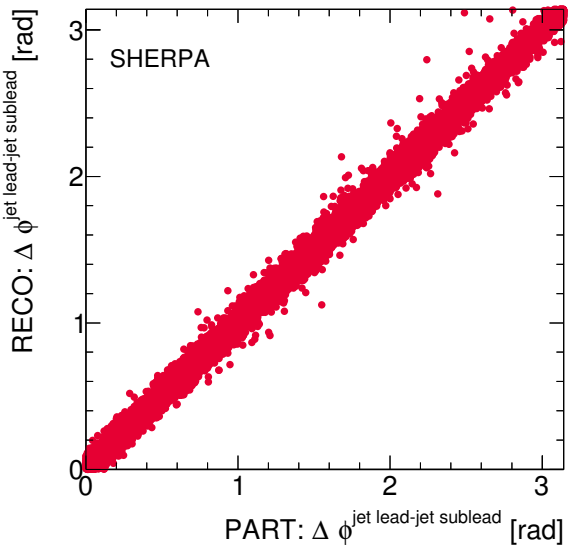


(a)

(b)

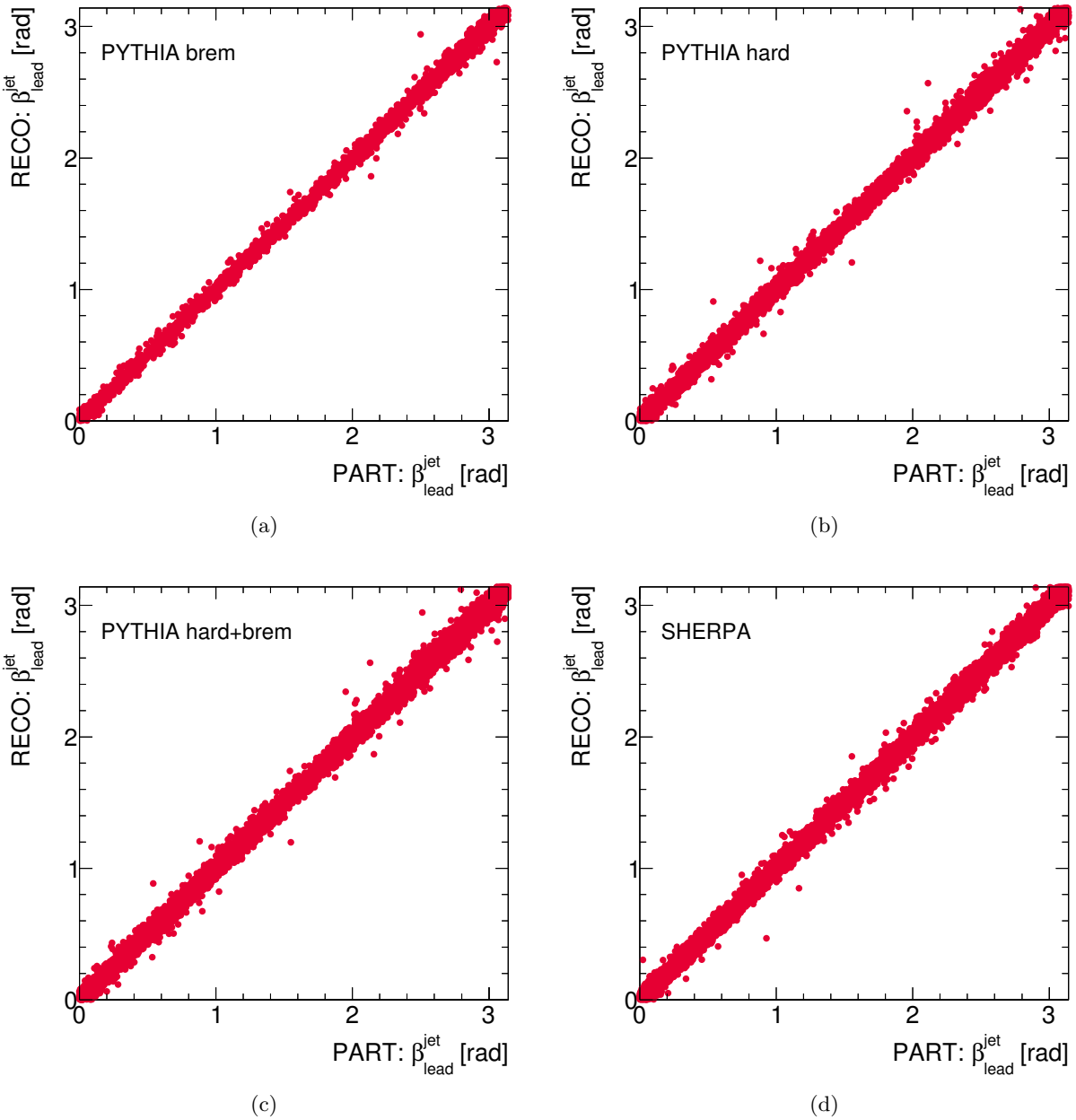


(c)



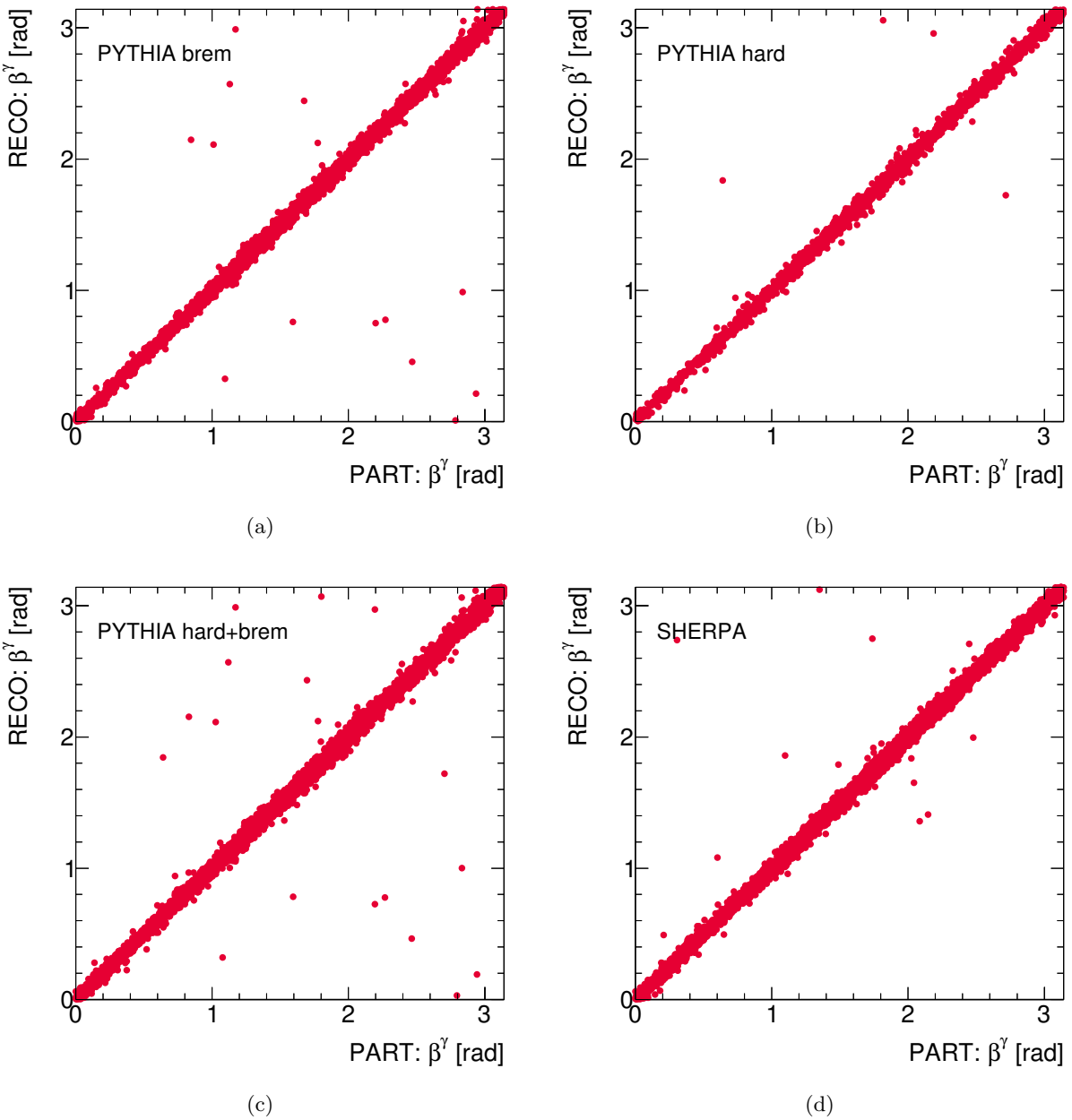
(d)

Figure 7.18: *The* $(\Delta\phi^{\text{jet lead}-\text{jet sublead}})_{\text{reco}}$ *vs.* $(\Delta\phi^{\text{jet lead}-\text{jet sublead}})_{\text{part}}$ *correlation for the photon+two-jets sample for* PYTHIA *brem (a),* PYTHIA *hard (b),* PYTHIA *(c) brem+hard and* SHERPA *(d).*



(a) (b) (c) (d)

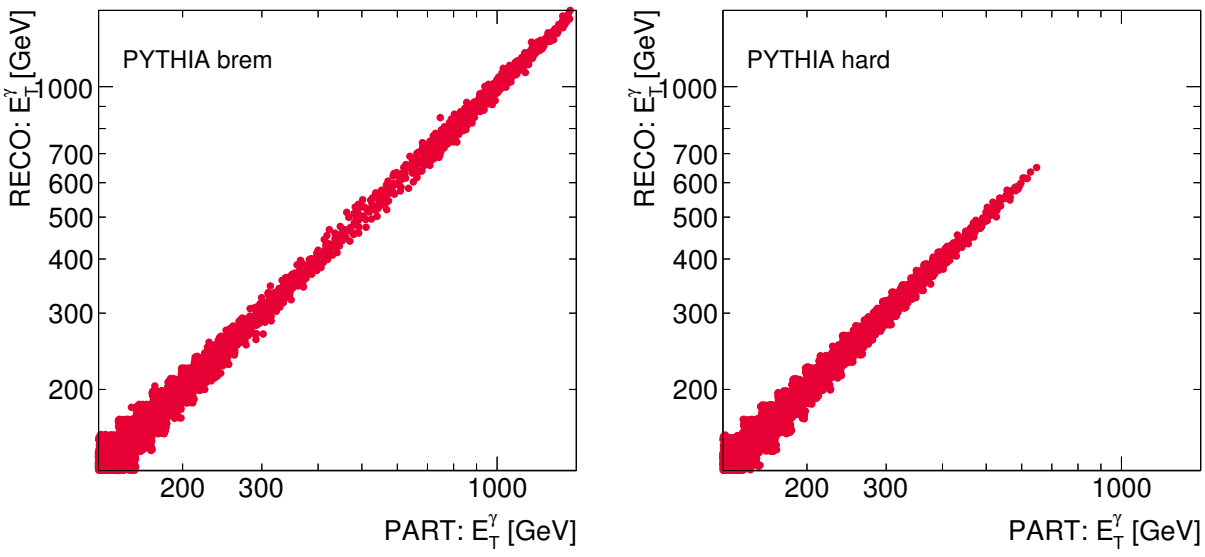
Figure 7.19: *The* $(\beta^{\text{jet}}_{\text{lead}})_{\text{reco}}$ *vs.* $(\beta^{\text{jet}}_{\text{lead}})_{\text{part}}$ *correlation for the colour coherence samples for* PYTHIA *brem (a),* PYTHIA *hard (b),* PYTHIA *(c) brem+hard and* SHERPA *(d).*



(a) (b)

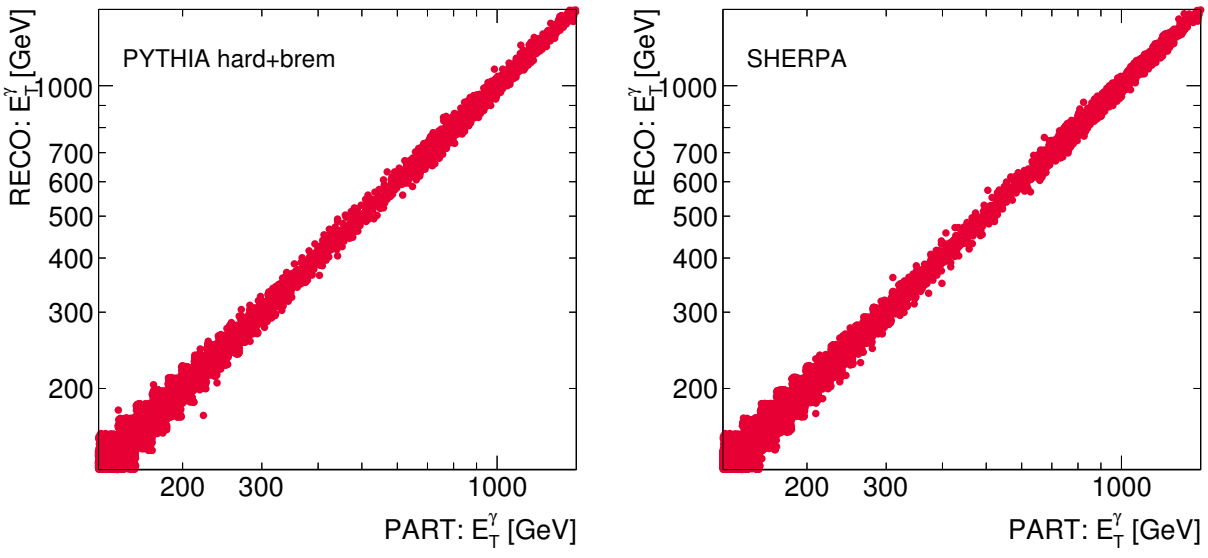
(c) (d)

Figure 7.20: *The* $(\beta^{\gamma})_{\mathrm{reco}}$ *vs.* $(\beta^{\gamma})_{\mathrm{part}}$ *correlation for the colour coherence samples for* PYTHIA *brem (a),* PYTHIA *hard (b),* PYTHIA *(c) brem+hard and* SHERPA *(d).*



(a)

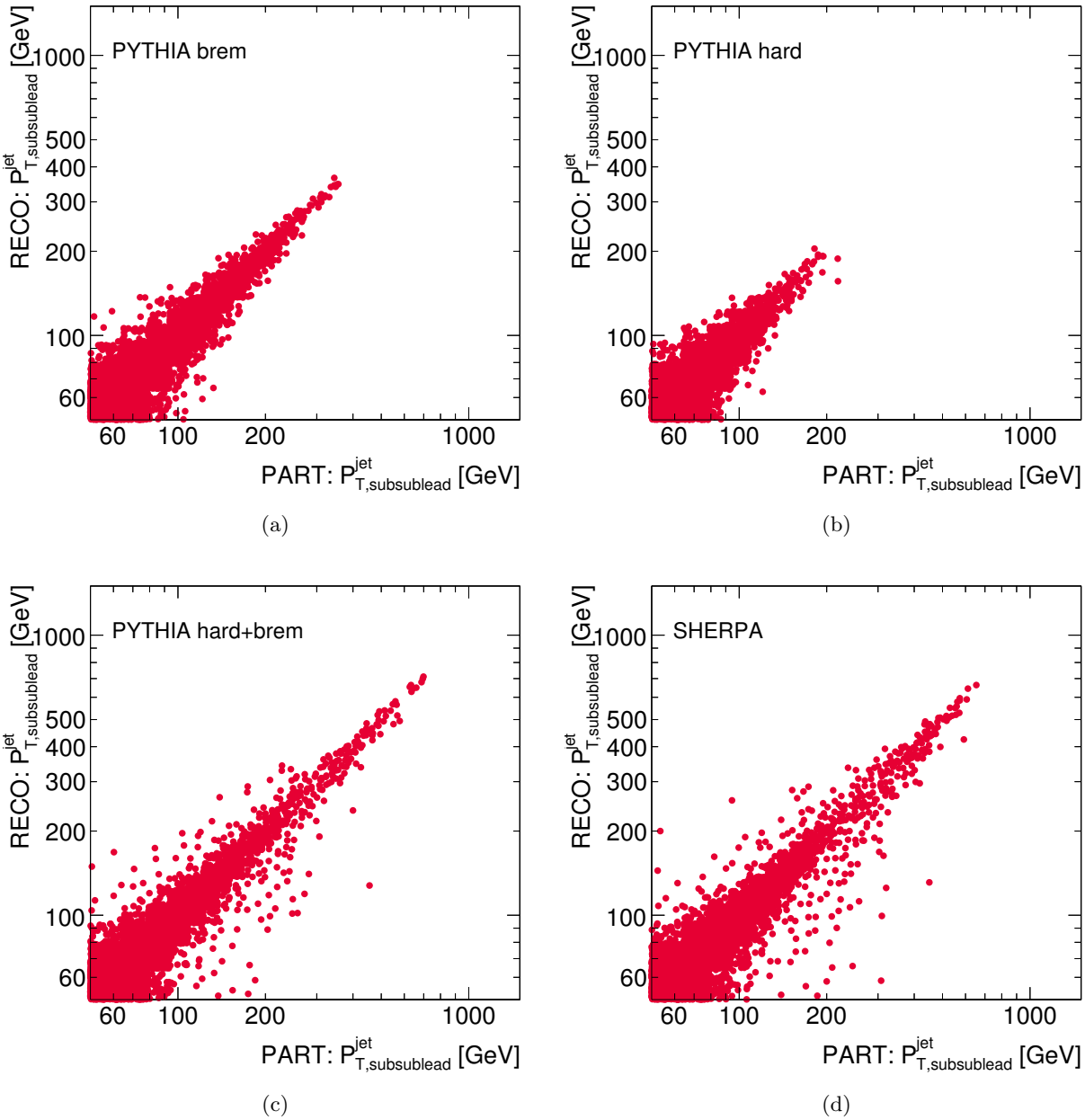
(b)



(c)

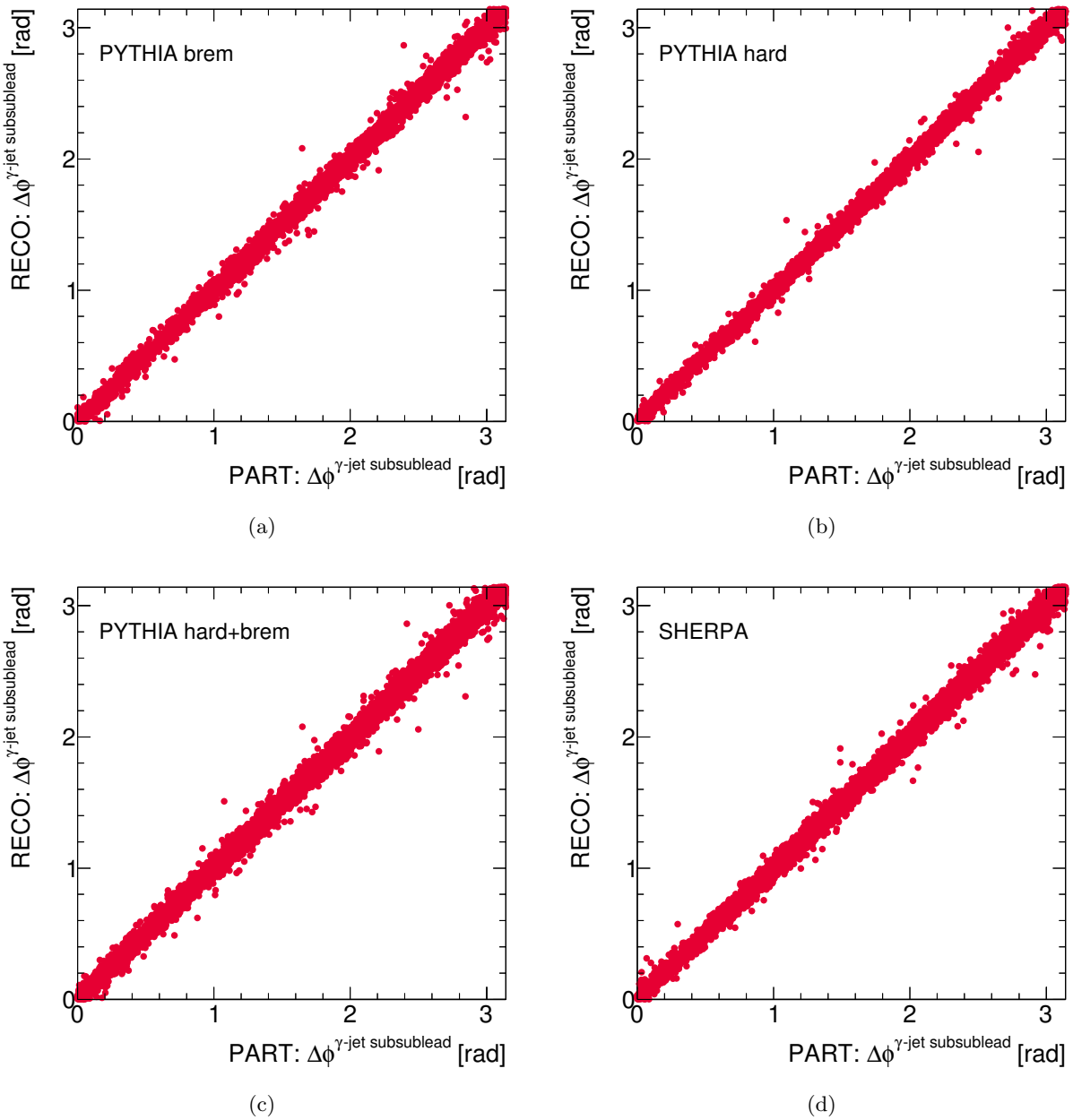
(d)

Figure 7.21: *The* $(E_{\mathrm{T}}^{\gamma})_{\mathrm{reco}}$ *vs.* $(E_{\mathrm{T}}^{\gamma})_{\mathrm{part}}$ *correlation for the photon+three-jets sample for* PYTHIA *brem (a),* PYTHIA *hard (b),* PYTHIA *(c) brem+hard and* SHERPA *(d).*



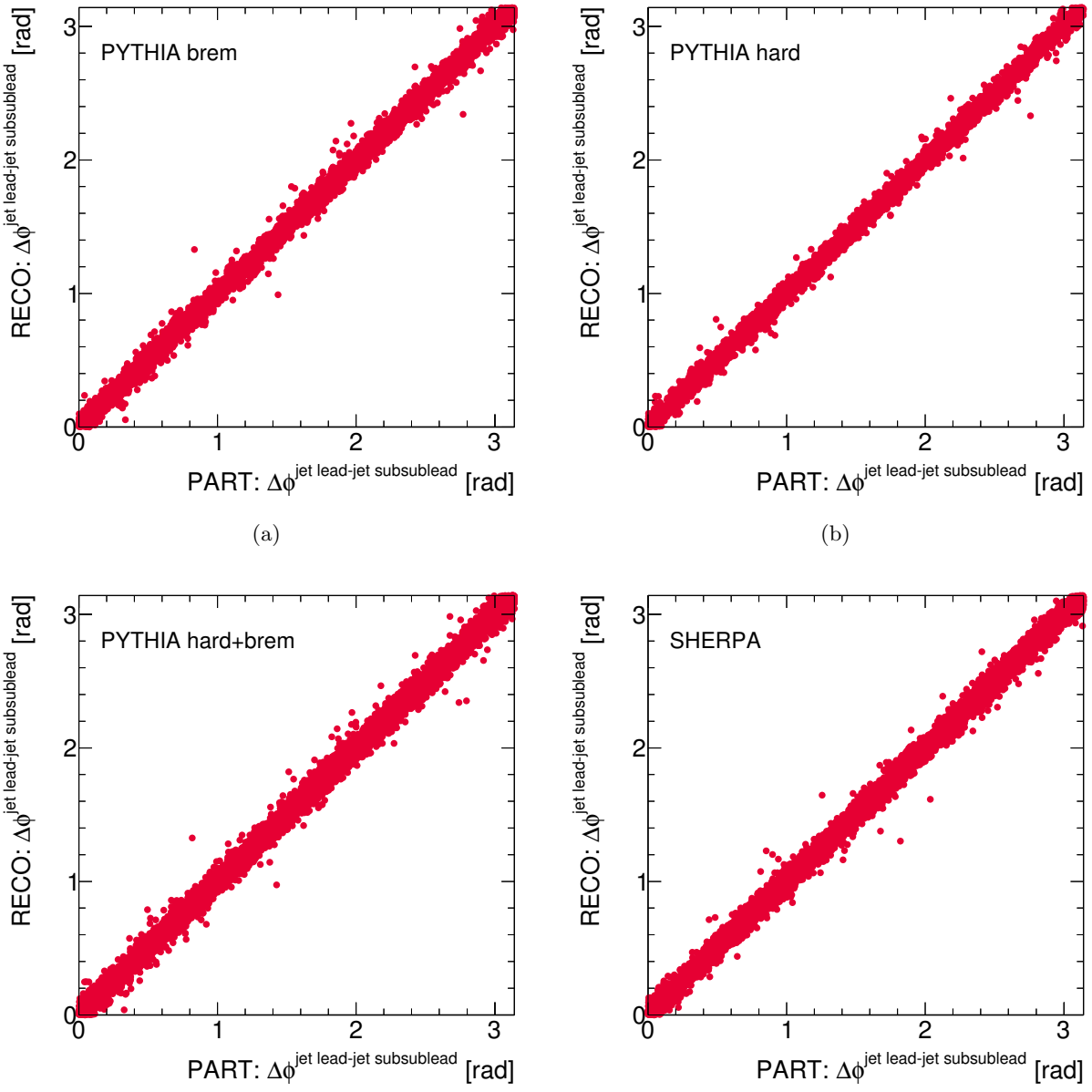
(a) (b) (c) (d)

Figure 7.22: *The* $(p^{\text{jet}}_{\text{T,subsublead}})_{\text{reco}}$ *vs.* $(p^{\text{jet}}_{\text{T,subsublead}})_{\text{part}}$ *correlation for the photon+three-jets sample for* PYTHIA *brem (a),* PYTHIA *hard (b),* PYTHIA *(c) brem+hard and* SHERPA *(d).*



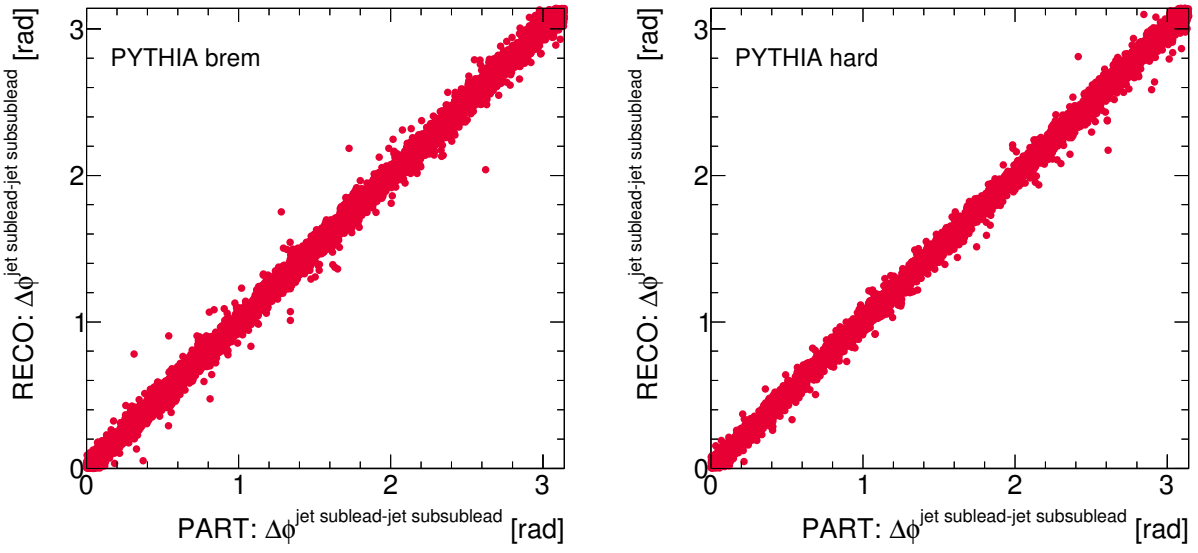
(a) (b) (c) (d)

Figure 7.23: *The* $(\Delta\phi^{\gamma-\text{jet subsublead}})_{\text{reco}}$ *vs.* $(\Delta\phi^{\gamma-\text{jet subsublead}})_{\text{part}}$ *correlation for the photon+three-jets sample for* PYTHIA *brem (a),* PYTHIA *hard (b),* PYTHIA *(c) brem+hard and* SHERPA *(d).*



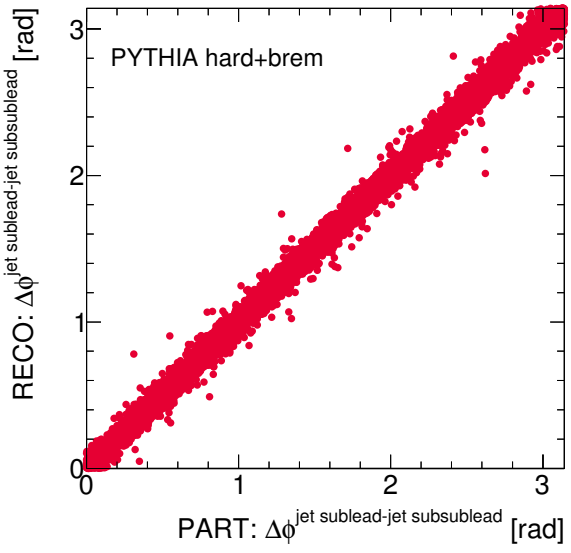
(a) (b) (c) (d)

Figure 7.24: *The* $(\Delta\phi^{\text{jet lead}-\text{jet subsublead}})_{\text{reco}}$ *vs.* $(\Delta\phi^{\text{jet lead}-\text{jet subsublead}})_{\text{part}}$ *correlation for the photon+three-jets sample for* PYTHIA *brem (a),* PYTHIA *hard (b),* PYTHIA *(c) brem+hard and* SHERPA *(d).*

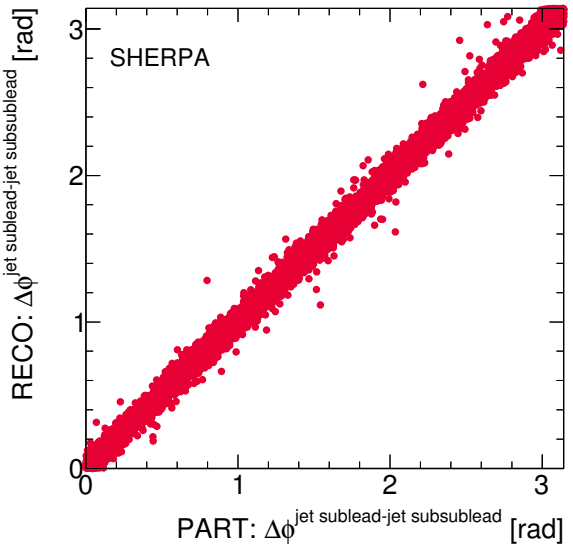


(a)

(b)



(c)



(d)

Figure 7.25: *The* $(\Delta\phi^{\text{jet sublead}-\text{jet subsublead}})_{\text{reco}}$ *vs.* $(\Delta\phi^{\text{jet sublead}-\text{jet subsublead}})_{\text{part}}$ *correlation for the photon+three-jets sample for* PYTHIA *brem (a),* PYTHIA *hard (b),* PYTHIA *(c) brem+hard and* SHERPA *(d).*

7.4 Cross-check using bayesian unfolding

A more sofisticated unfolding method, based on the iterative application of Bayes' theorem [96], was investigated to cross-check and validate the results obtained using the bin-by-bin method explained in section 6.10. It was implemented using the RooUnfold package [97].

The bayesian unfolding method takes properly into account the migrations between bins and the purity and efficiency of the selection criteria when the MC description of the data is not adequate.

The main requirement of this method is to have sufficiently large MC samples to construct the reference matrices, since otherwise, the results are less reliable than those obtained from the bin-by-bin method because they are more affected by such lack of statistics. Another problem is the abnormal blow up of the statistical uncertainty of the resulting cross sections which appears when a large number of iterations after convergence is used. To make the comparison between the bayesian and the bin-by-bin methods more meaningful, these issues were investigated before attempting the cross-check.

Figures 7.26 to 7.28 show the relative difference between the results obtained based on N iterations with respect to those using $N+1$ iterations as a function of the number of iterations for the observables studied; the comparison was done using the sum of the absolute value of the relative differences of each bin of each cross section. The relative difference between iterations 1 and 2 is quite large; around iteration 4, the differences start to be very small for all observables. Therefore, the nominal bayesian unfolding was performed using 4 iterations. Figures 7.29 to 7.31 show the relative difference between the results obtained using N iterations with respect to those obtained using 4 iterations for the observables studied. It is observed that the oscillations on the results for $N < 4$ are quite large, these oscillations disappear for $N > 4$.

The comparison between the cross sections unfolded via the bin-by-bin and the bayesian, using $N = 4$, methods for all observables is shown in figures 7.32 to 7.34, using the nominal MC for each observable, Pythia or Sherpa. The differences between the cross sections obtained via the two methods are smaller than 1%, except for the first bin for each $p_{\mathrm{T}}^{\mathrm{jet}}$ distributions (due to threshold effects) and in the regions of phase space where the MC statistics is poor. For comparison, the statistical uncertainty of the cross sections is also included in the figures and shows that the size of this uncertainty is typically of the same size or bigger than the difference of the cross sections obtained with the two methods. The correlations matrices from Pythia or Sherpa used in the bayesian unfolding are shown in figures 7.35 to 7.37.

All the results shown validate the use of the bin-by-bin unfolding as the nominal method to measure the cross sections.

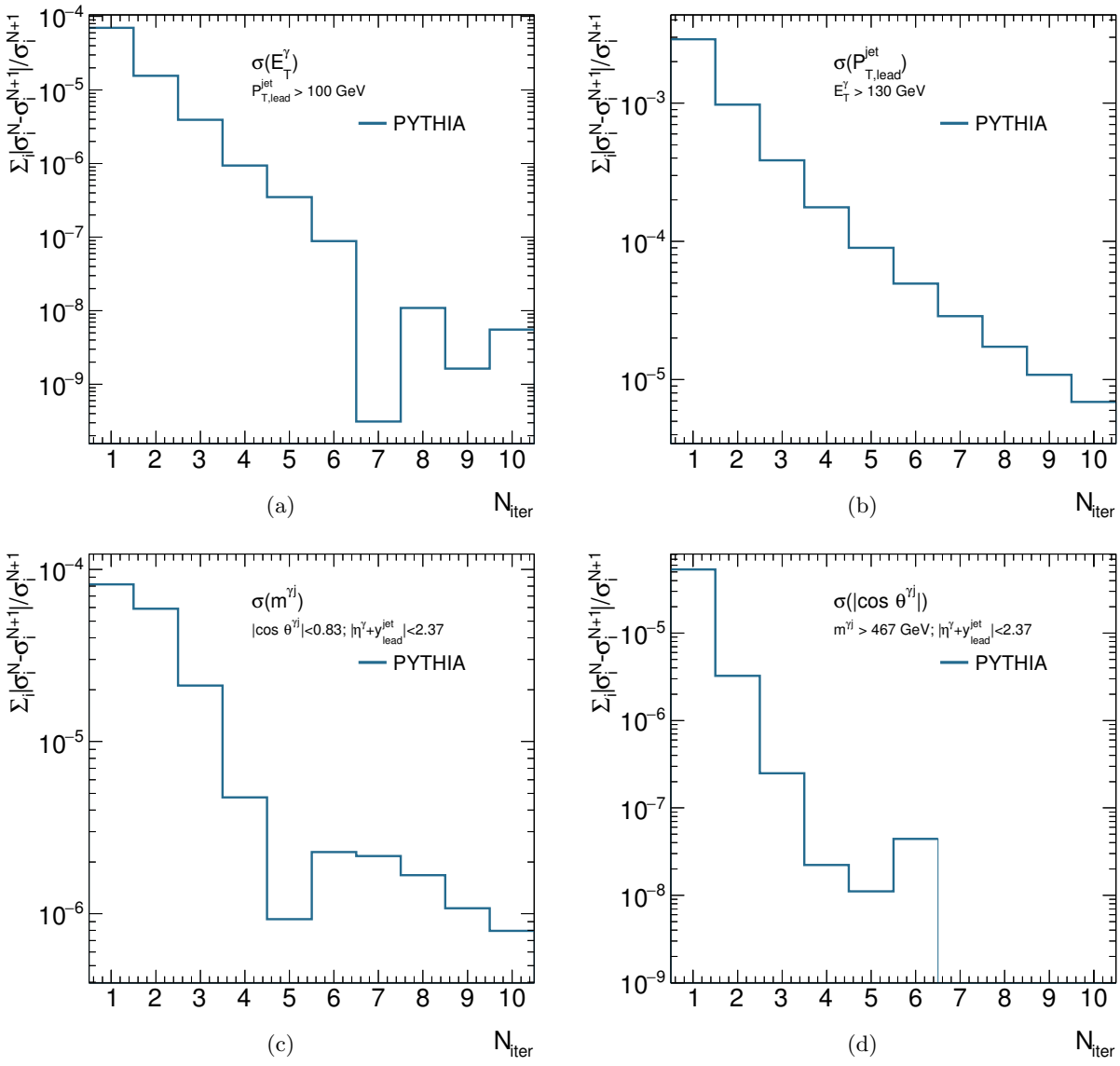


Figure 7.26: *Relative differences between the cross sections unfolded using N and $N+1$ iterations as functions of the number of iterations for E_{T}^{γ} (a), $p_{\mathrm{T,lead}}^{\mathrm{jet}}$ (b), $m^{\gamma j}$ (c) and $|\cos\theta^{\gamma j}|$ (d) for the photon+1-jet sample.*

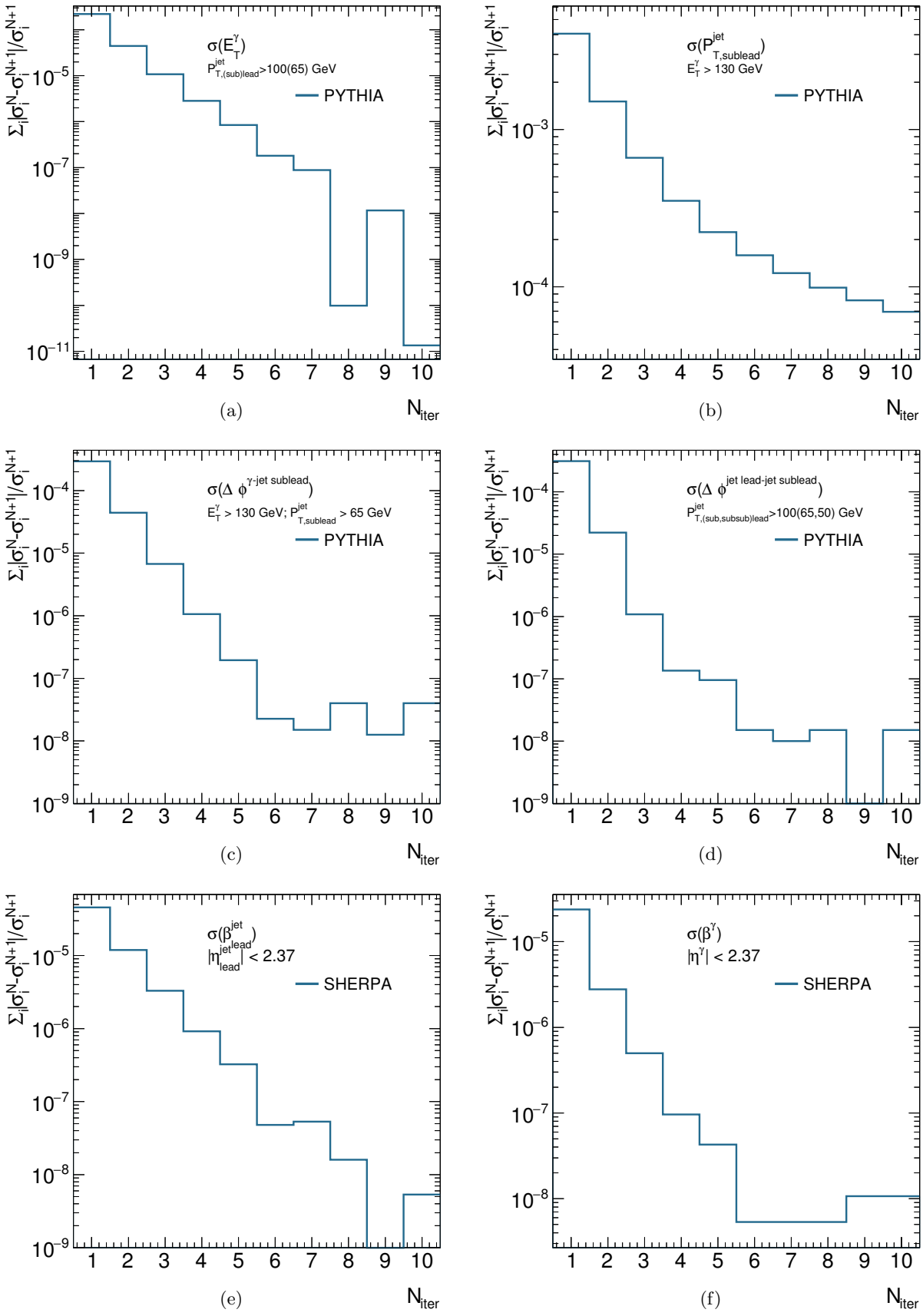


Figure 7.27: *Relative differences between the cross sections unfolded using N and $N+1$ iterations as functions of the number of iterations for E_{T}^{γ} (a), $p_{\mathrm{T,sublead}}^{\mathrm{jet}}$ (b), $\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$ (c), $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$ (d), $\beta_{\mathrm{lead}}^{\mathrm{jet}}$ (e) and β^{γ} (f) for the photon+two-jets (a,b,c,d) and the colour coherence (e,f) samples.*

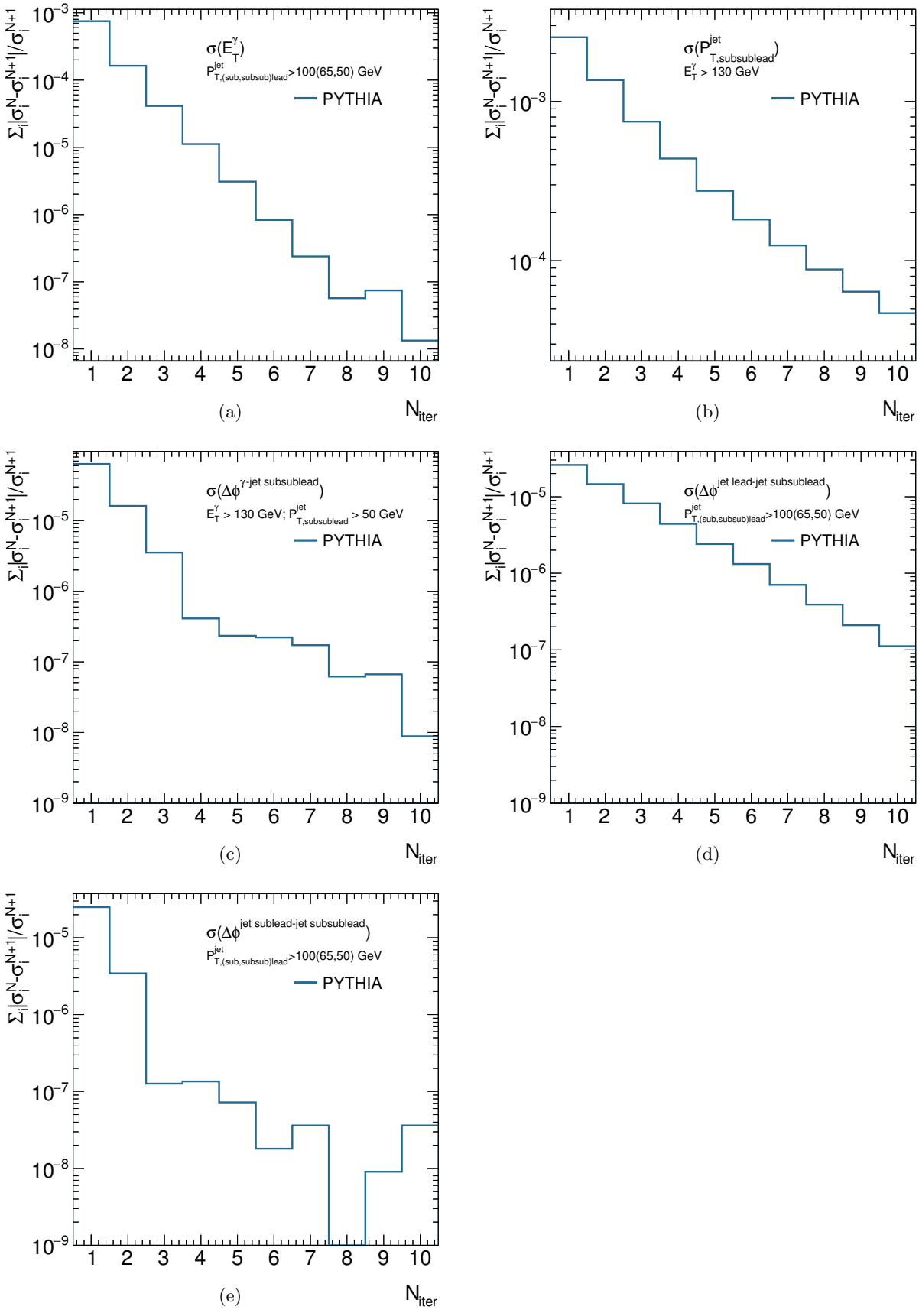


Figure 7.28: *Relative differences between the cross sections unfolded using N and $N+1$ iterations as functions of the number of iterations for E_{T}^{γ} (a), $p_{\mathrm{T,subsublead}}^{\mathrm{jet}}$ (b), $\Delta\phi^{\gamma-\mathrm{jet\ subsublead}}$ (c), $\Delta\phi^{\mathrm{jet\ lead}-\mathrm{jet\ subsublead}}$ (d), $\Delta\phi^{\mathrm{jet\ sublead}-\mathrm{jet\ subsublead}}$ (e) for the photon+three-jets sample.*

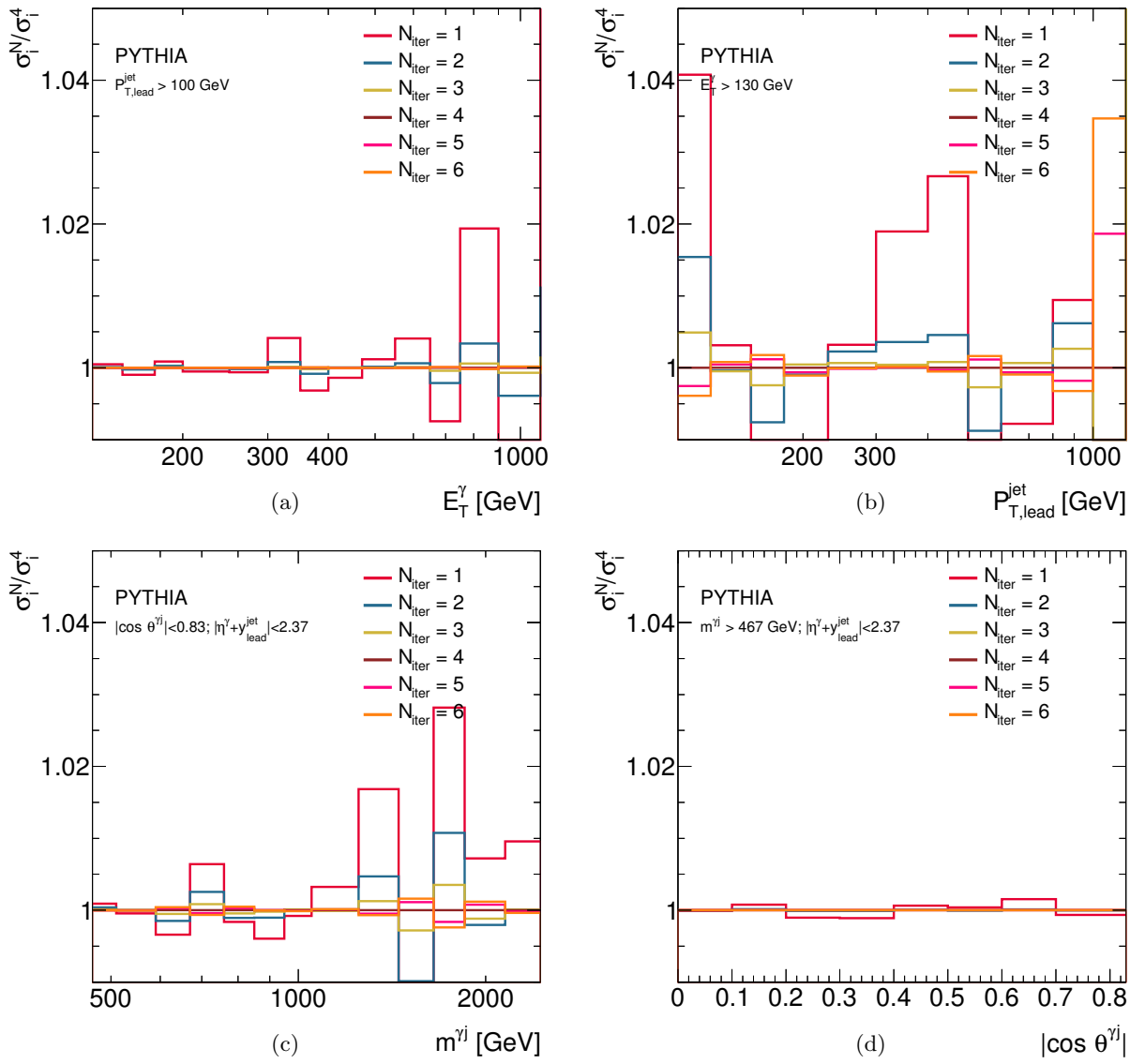


Figure 7.29: *Ratio between the cross sections unfolded using the bayesian method with N iterations with respect to the results unfolded using 4 iterations as a function of* E^{γ}_{T} *(a),* $p^{\mathrm{jet}}_{\mathrm{T,lead}}$ *(b),* $m^{\gamma j}$ *(c) and* $|\cos\theta^{\gamma j}|$ *(d) for the photon+1-jet sample.*

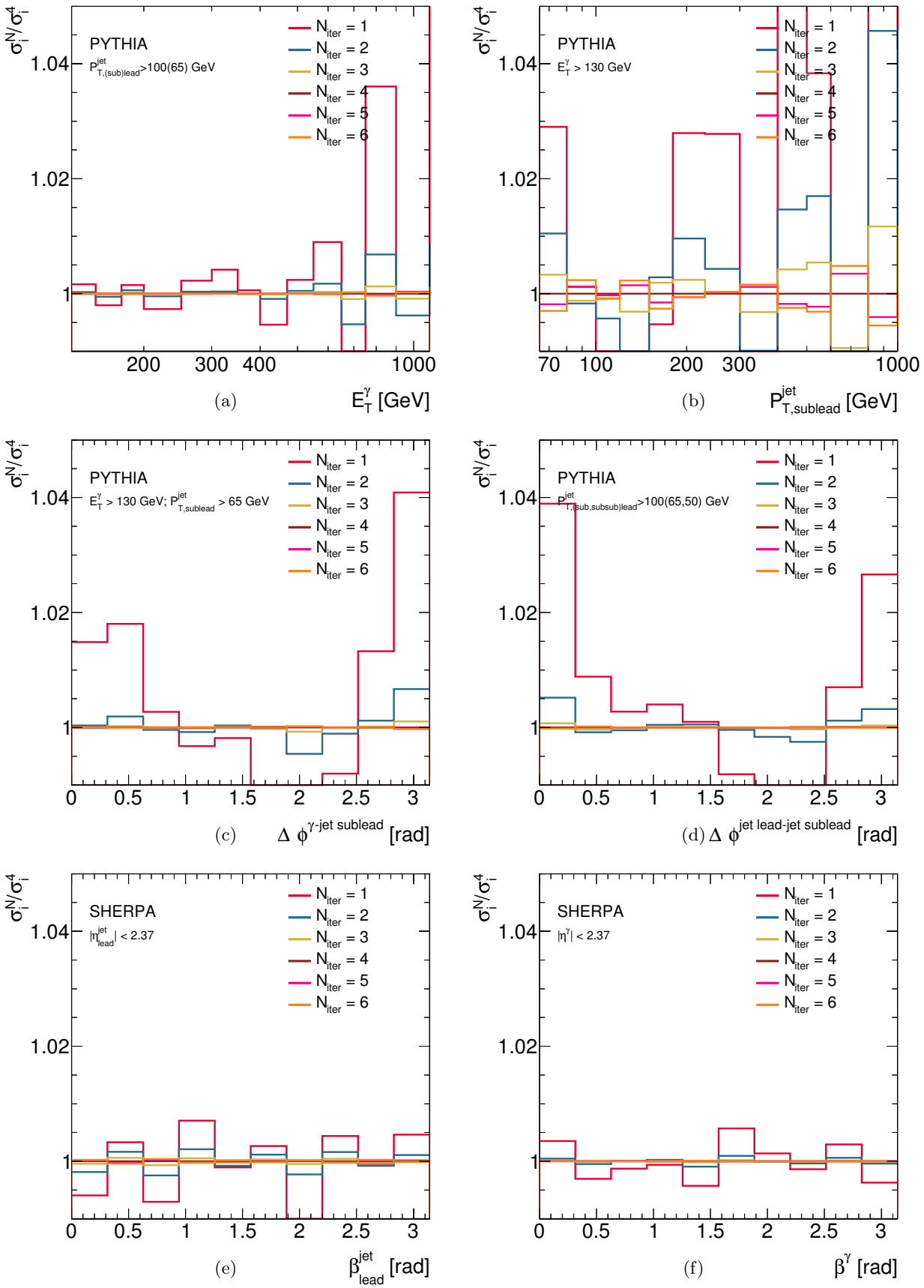


Figure 7.30: *Ratio between the cross sections unfolded using the bayesian method with N iterations with respect to the results unfolded using 4 iterations as a function of* E^γ_T *(a),* $p^{jet}_{T,sublead}$ *(b),* $\Delta\phi^{\gamma-\text{jet sublead}}$ *(c),* $\Delta\phi^{\text{jet lead}-\text{jet sublead}}$ *(d),* β^{jet}_{lead} *(e) and* β^γ *(f) for the photon+two-jets (a,b,c,d) and the colour coherence (e,f) samples.*

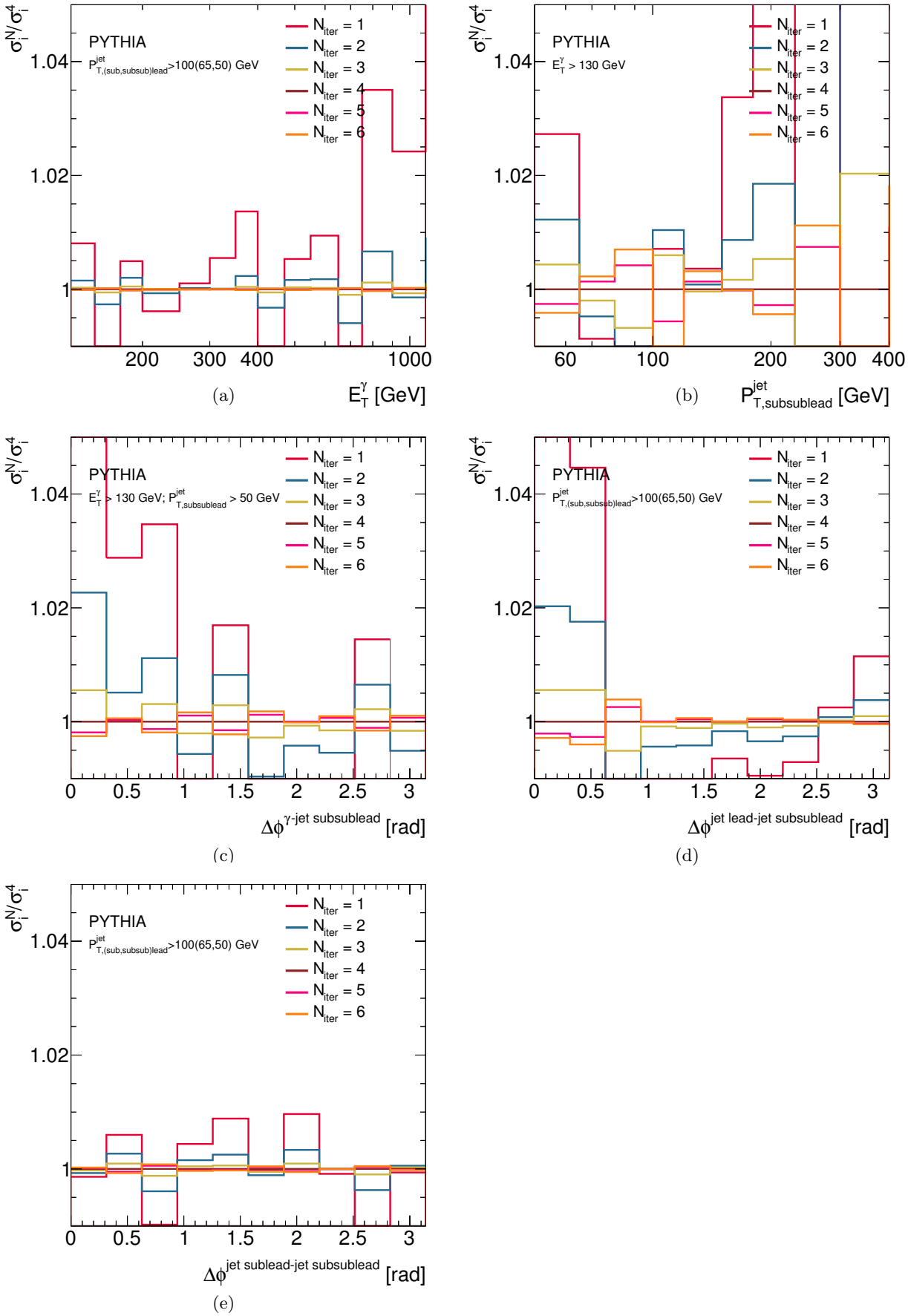


Figure 7.31: *Ratio between the cross sections unfolded using the bayesian method with N iterations with respect to the results unfolded using 4 iterations as a function of* E^{γ}_{T} *(a),* $p^{\mathrm{jet}}_{\mathrm{T,subsublead}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\ subsublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\ lead}-\mathrm{jet\ subsublead}}$ *(d),* $\Delta\phi^{\mathrm{jet\ sublead}-\mathrm{jet\ subsublead}}$ *(e) for the photon+three-jets sample.*

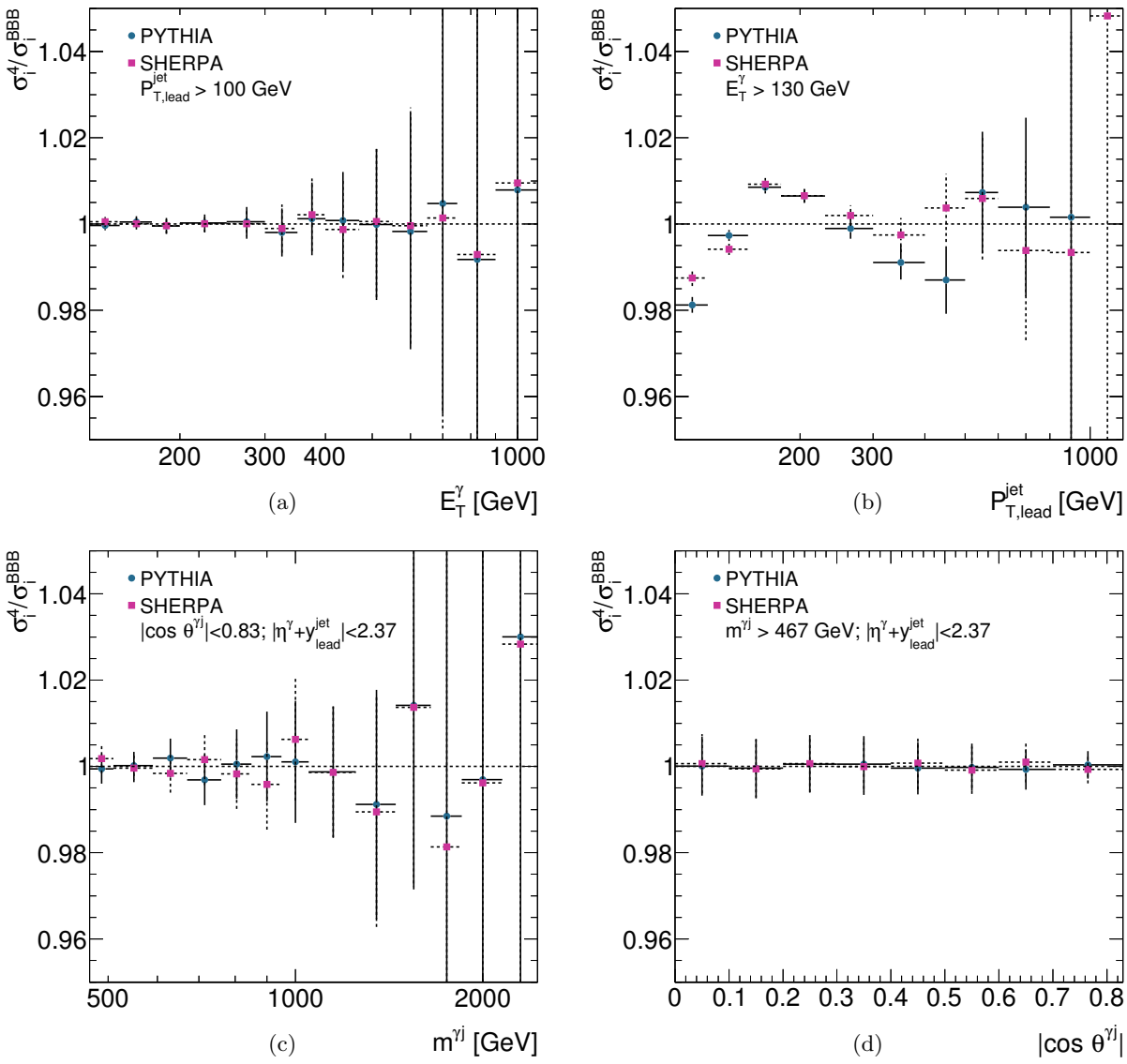


Figure 7.32: *Ratio between the cross sections unfolded using the bayesian method with 4 iterations and the nominal cross sections as a function of E_{T}^{γ} (a), $p_{\mathrm{T,lead}}^{\mathrm{jet}}$ (b), $m^{\gamma j}$ (c) and $|\cos\theta^{\gamma j}|$ (d) for the photon+1-jet sample.*

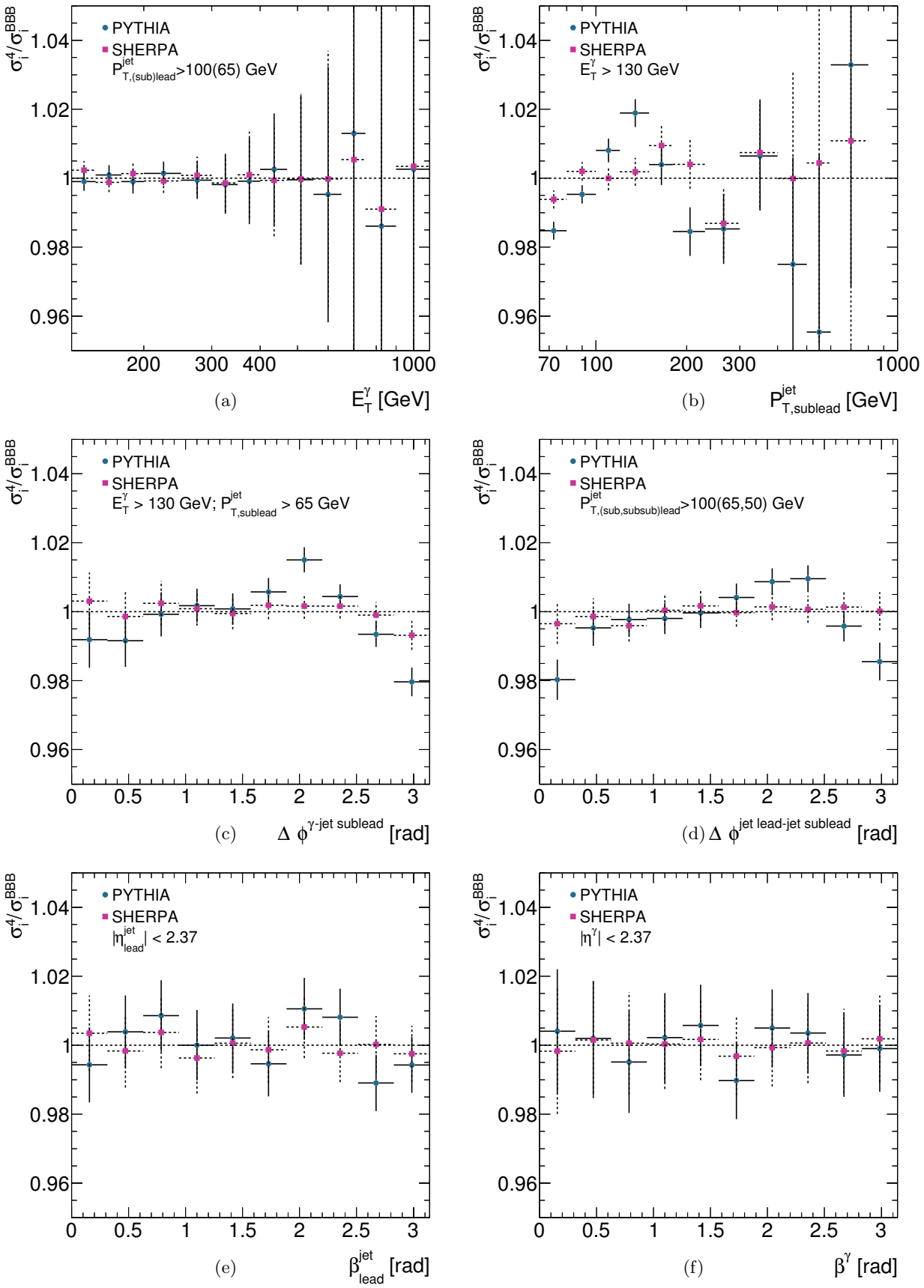


Figure 7.33: *Ratio between the cross sections unfolded using the bayesian method with 4 iterations and the nominal cross sections as a function of* E_{T}^{γ} *(a),* $p_{\mathrm{T,sublead}}^{\mathrm{jet}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$ *(d),* $\beta_{\mathrm{lead}}^{\mathrm{jet}}$ *(e) and* β^{γ} *(f) for the photon+two-jets (a,b,c,d) and the colour coherence (e,f) samples.*

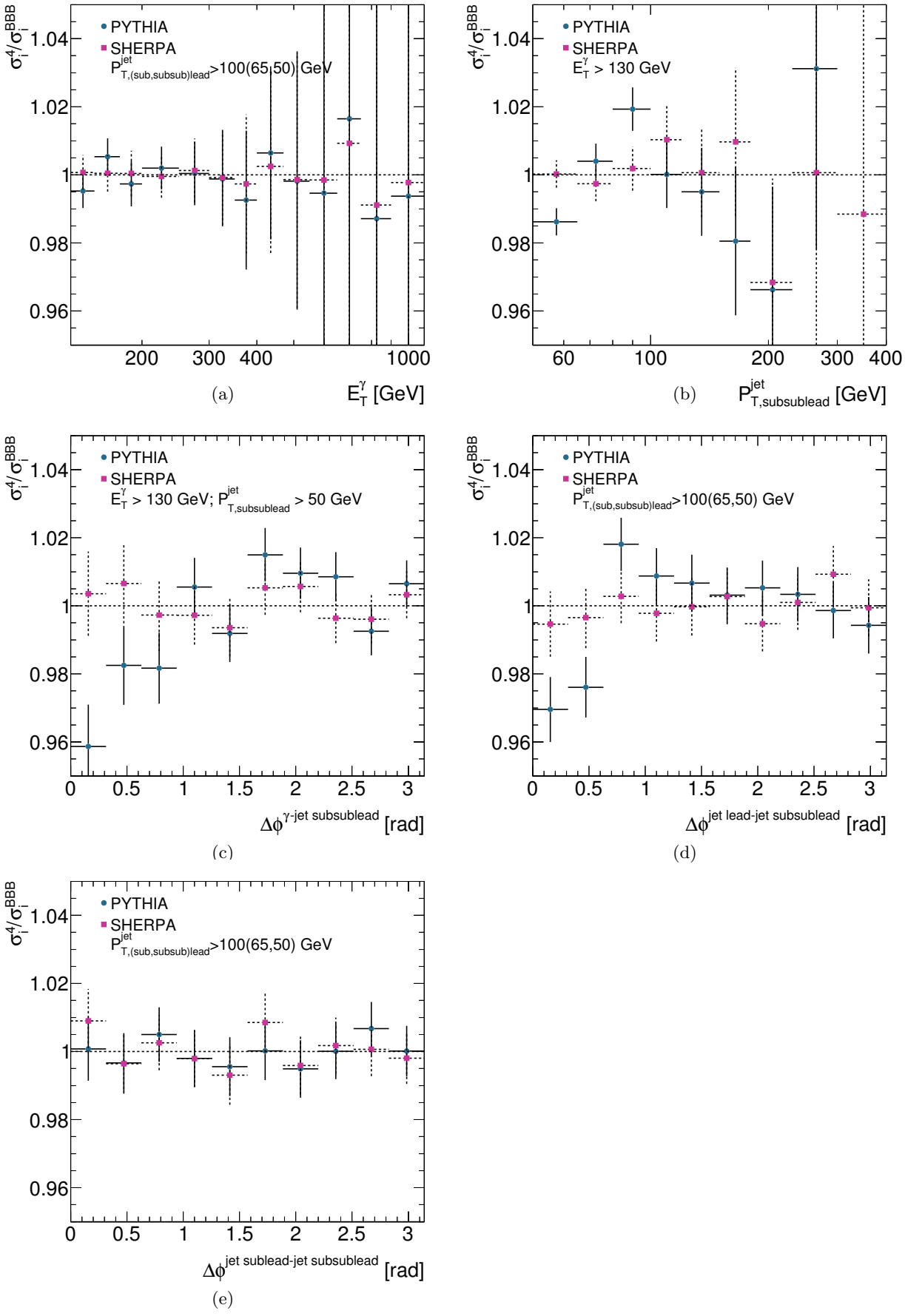


Figure 7.34: *Ratio between the cross sections unfolded using the bayesian method with 4 iterations and the nominal cross sections as a function of* E_{T}^{γ} *(a),* $p_{\mathrm{T,subsublead}}^{\mathrm{jet}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\ subsublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\ lead}-\mathrm{jet\ subsublead}}$ *(d),* $\Delta\phi^{\mathrm{jet\ sublead}-\mathrm{jet\ subsublead}}$ *(e) for the photon+three-jets sample.*

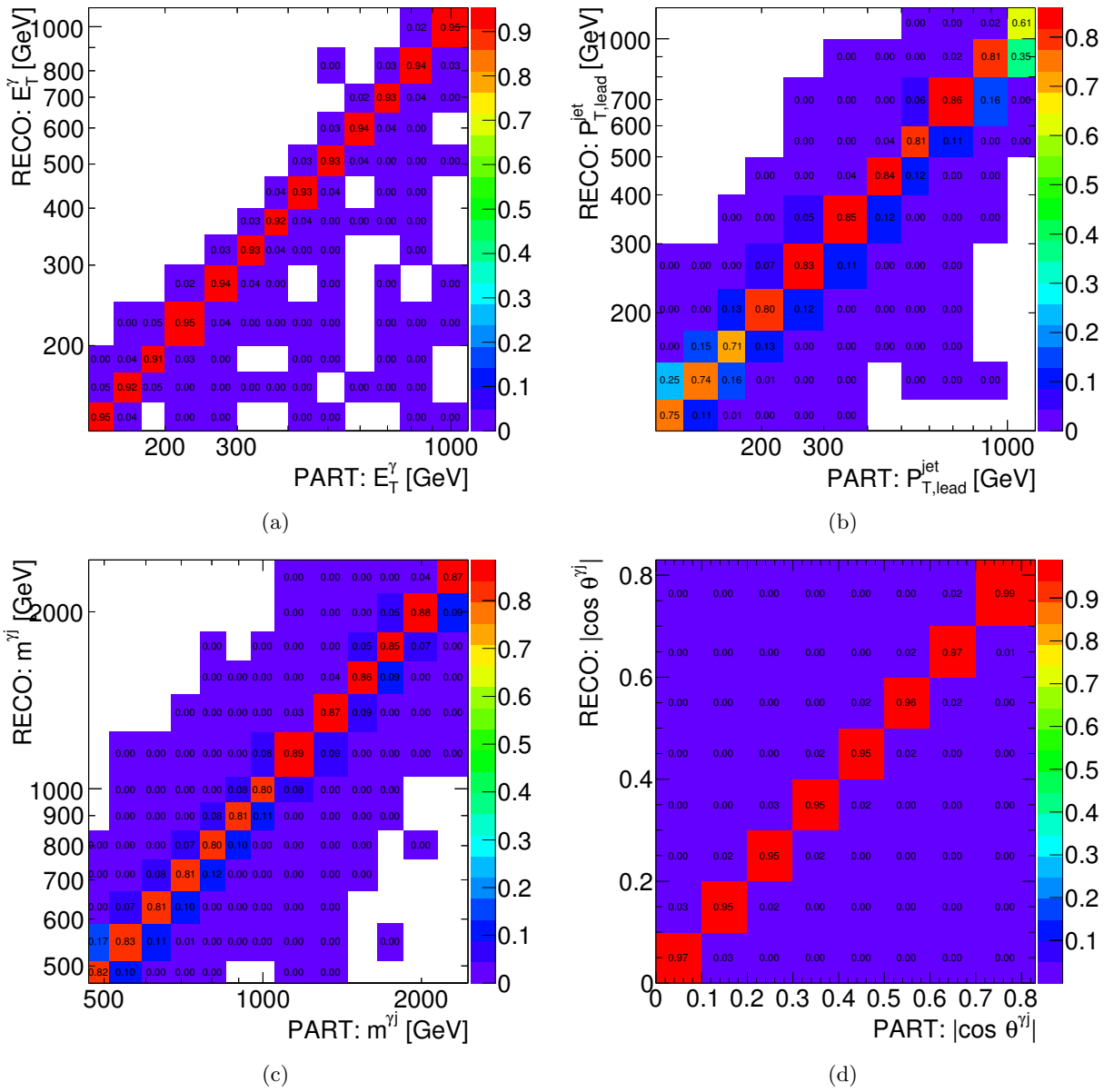


Figure 7.35: *Correlation matrices for E_{T}^{γ} (a), $p_{\mathrm{T,lead}}^{\mathrm{jet}}$ (b), $m^{\gamma j}$ (c) and $|\cos\theta^{\gamma j}|$ (d) for the photon+1-jet sample.*

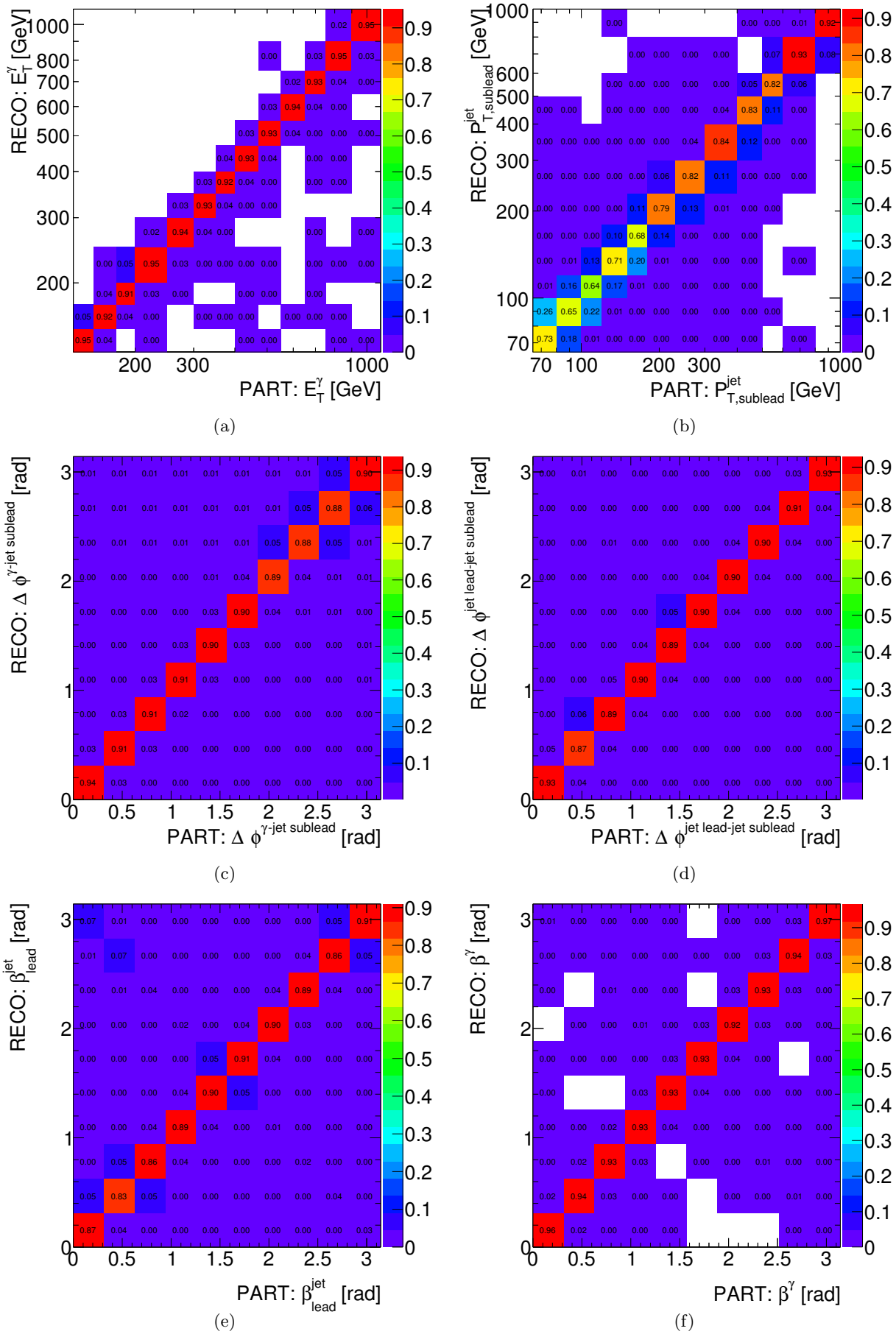
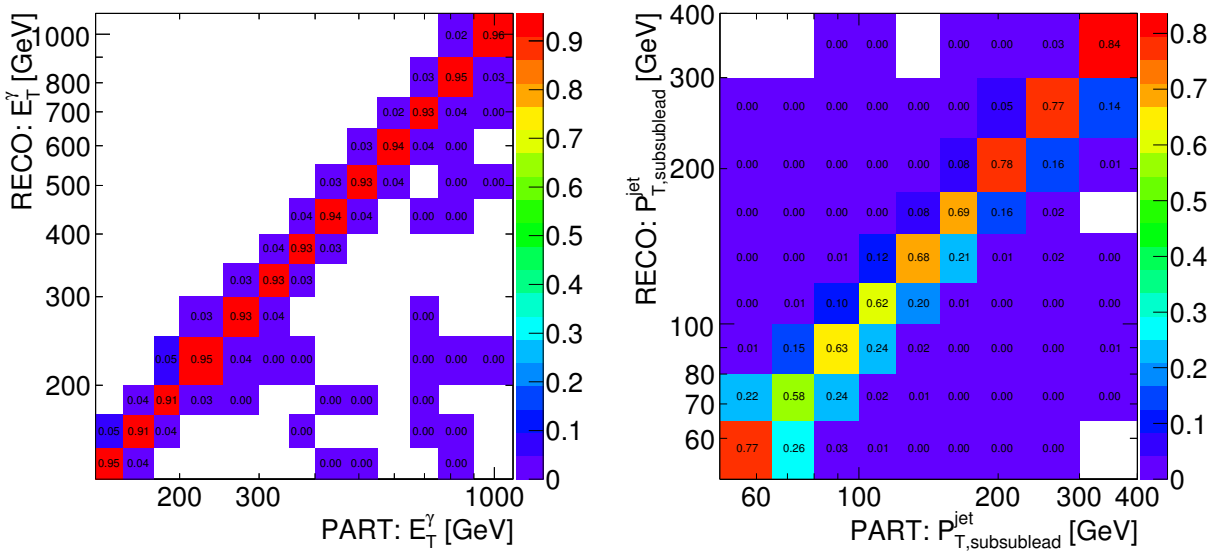
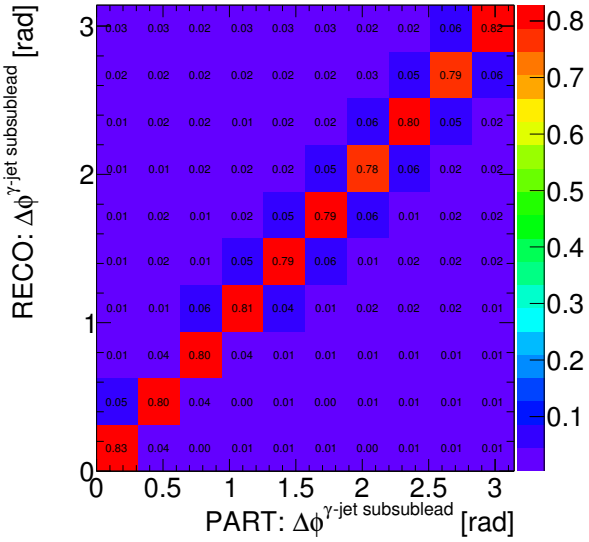


Figure 7.36: *Correlation matrices for E_T^γ (a), $p_{T,sublead}^{jet}$ (b), $\Delta\phi^{\gamma-\text{jet sublead}}$ (c), $\Delta\phi^{\text{jet lead}-\text{jet sublead}}$ (d), β_{lead}^{jet} (e) and β^γ (f) for the photon+two-jets (a,b,c,d) and the colour coherence (e,f) samples.*

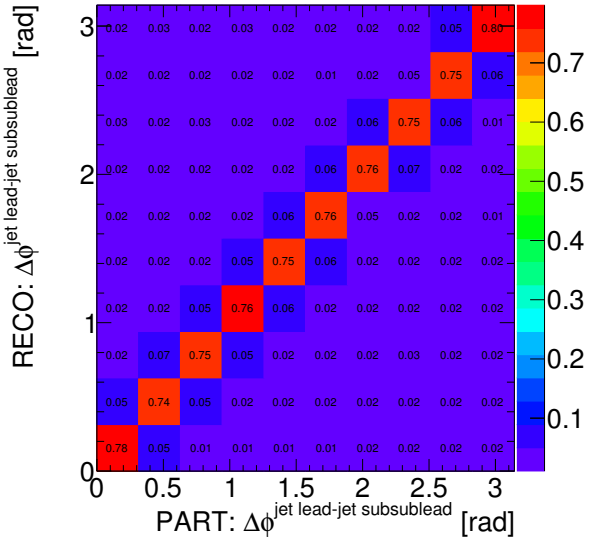


(a)

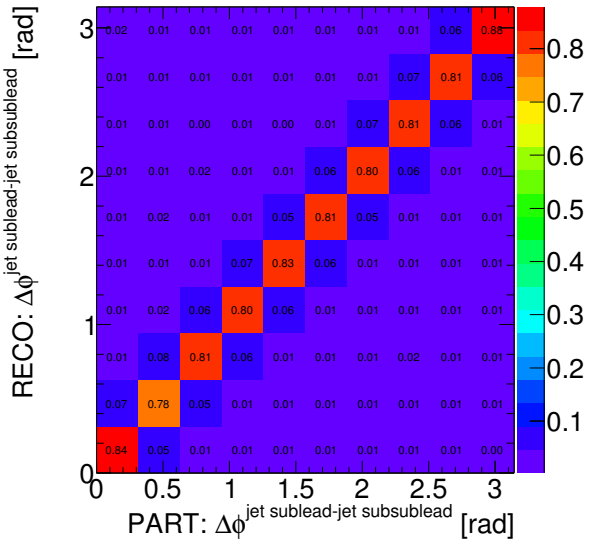
(b)



(c)



(d)



(e)

Figure 7.37: *Correlation matrices for* E_{T}^{γ} *(a),* $p_{\mathrm{T,subsublead}}^{\mathrm{jet}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\ subsublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\ lead}-\mathrm{jet\ subsublead}}$ *(d),* $\Delta\phi^{\mathrm{jet\ sublead}-\mathrm{jet\ subsublead}}$ *(e) for the photon+three-jets sample.*

7.5 Pile up studies

To investigate the effect of the pile-up on the measurements, several tests were performed and are presented below.

Differences between data and MC in the high and low pile-up regions can lead to discrepancies in the cross sections. This effect is expected to be negligible since the μ distribution in the MC was reweighted to describe the one in the data (see section 6.5). Still, some effects can remain since the correction was performed for the full sample. This was tested by separating the data and MC samples into low- ($\mu < 19$) and high- ($\mu > 19$) μ subsamples and the distributions compared. Figures 7.38 to 7.40 show the ratios of the high- and low-μ subsamples for data and MC together with the double ratio (lower part of the plots) as functions of the observables studied. The double ratios fluctuate around 1, which means that no remaining effect due to the different pile-up conditions in data and MC affects the measurements.

The effect of the different pile-up conditions on the signal purity was tested by computing the purity for low- and high-μ subsamples, which is estimated using a data-driven technique (see section 6.6). The comparison between these purities and the nominal is shown in Figures 7.41 to 7.43 as functions of the observables studied. No significant effect is observed. Also, the effect on the acceptance correction factors (which relies purily on the MC, see Section 6.10) was tested. Figures 7.44 to 7.46 show the correction factors for the nominal, low-μ and high-μ samples. Again, no significant effect is observed.

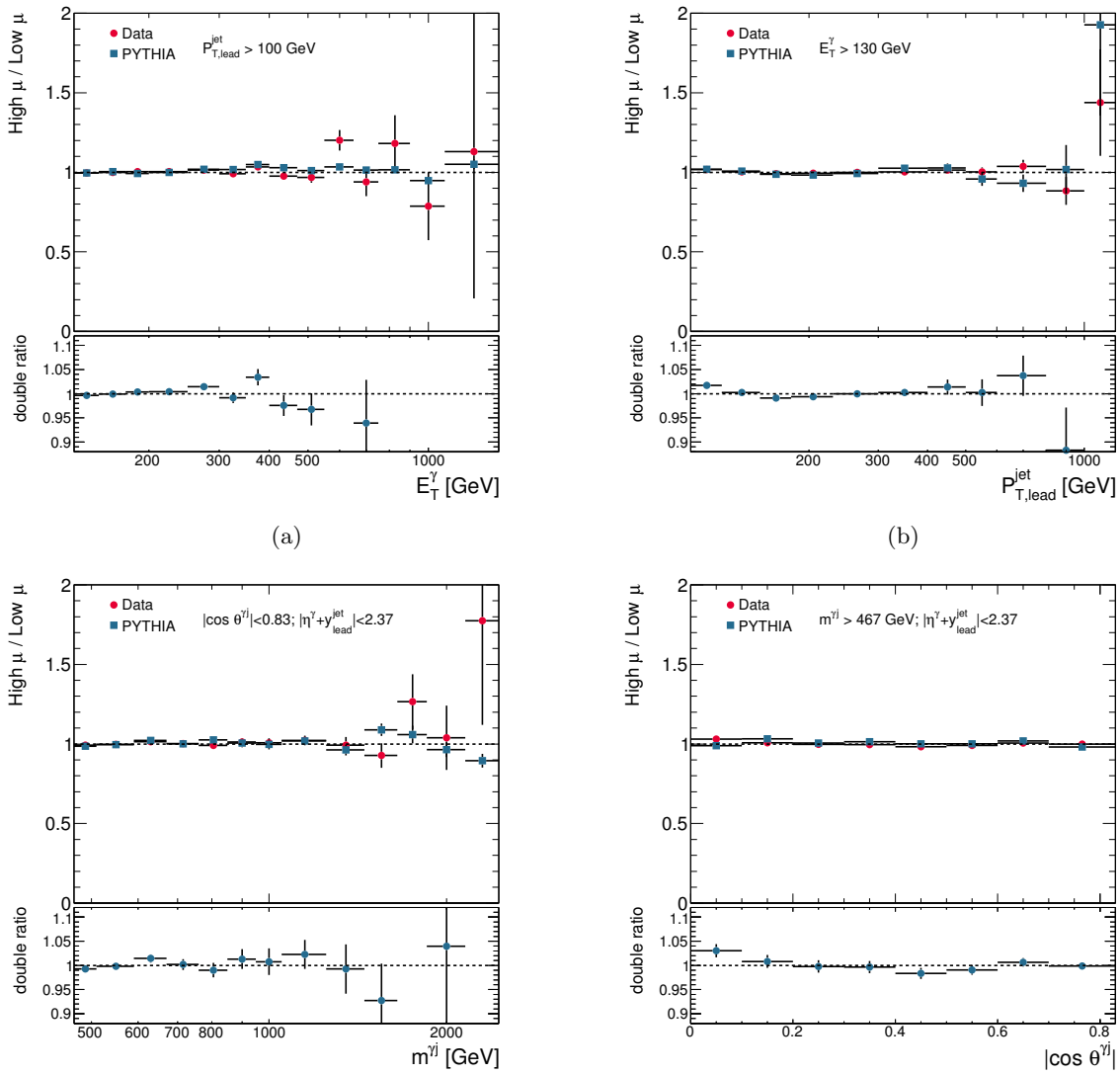


Figure 7.38: *Ratio of high and low μ in data (dots) and MC (squares) as a function of* E^{γ}_{T} *(a),* $p^{\mathrm{jet}}_{\mathrm{T,lead}}$ *(b),* $m^{\gamma j}$ *(c) and* $|\cos\theta^{\gamma j}|$ *(d) for the photon+1-jet sample. The double ratio is shown in the lower part of the figure.*

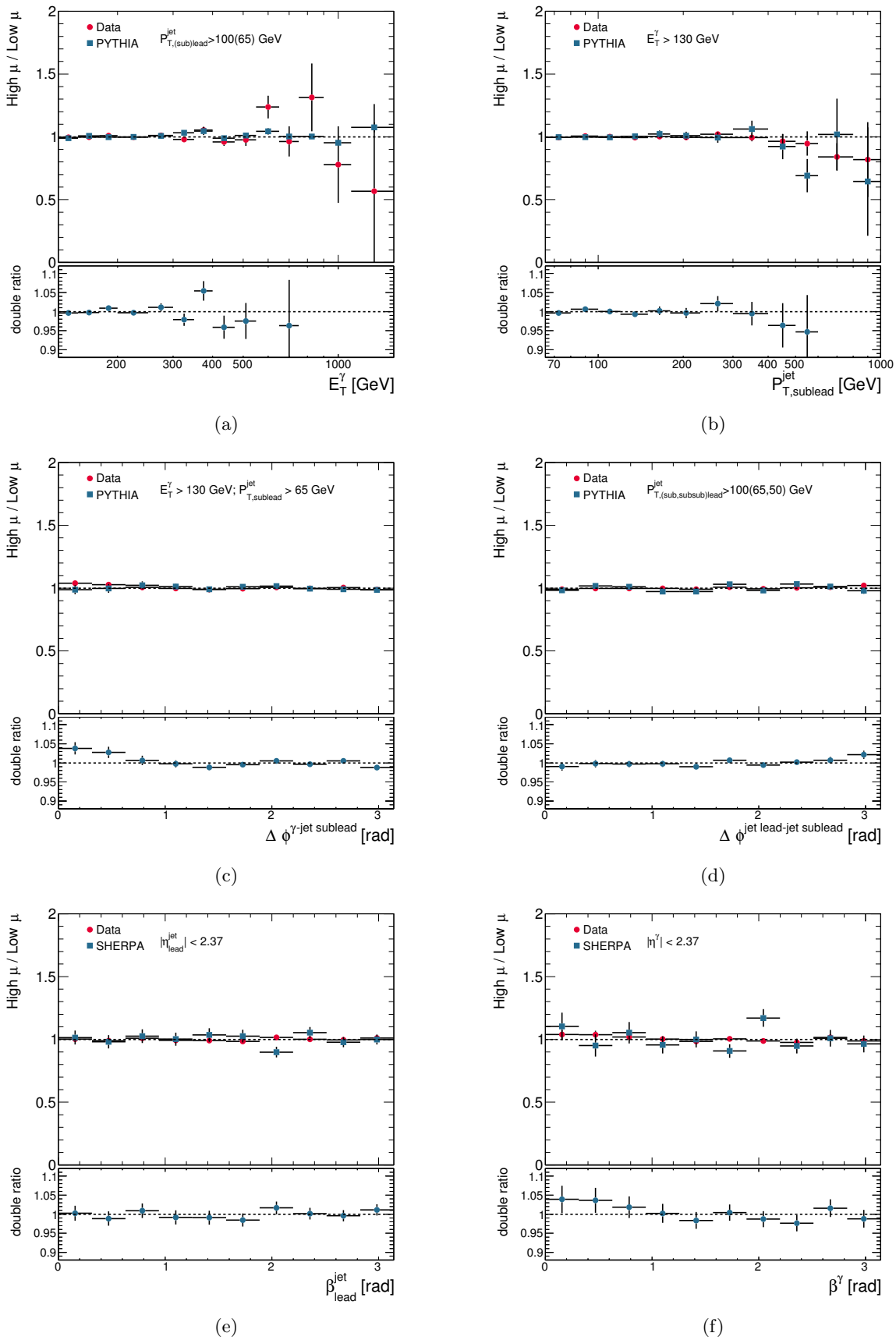
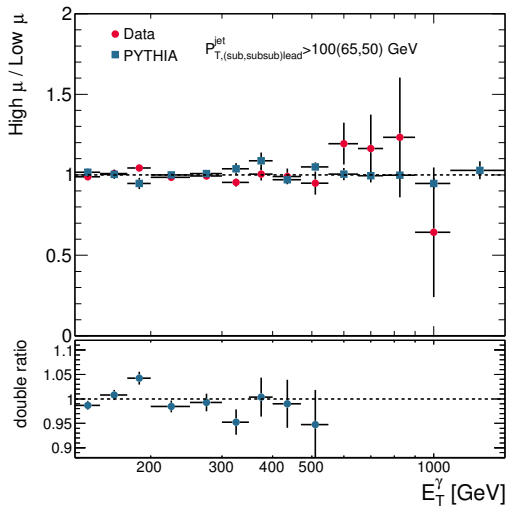
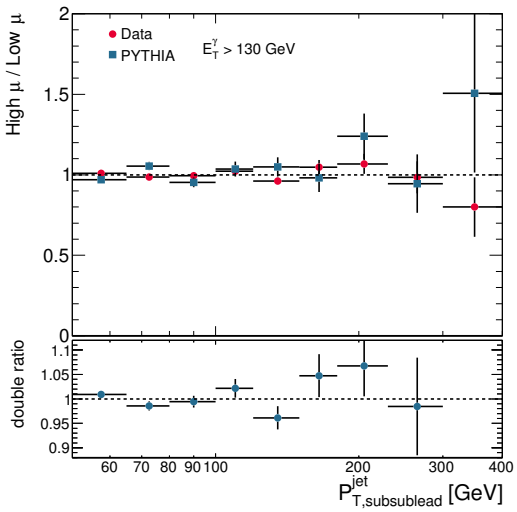


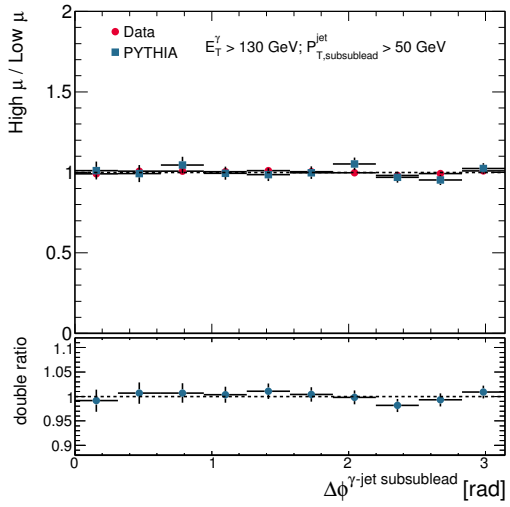
Figure 7.39: *Ratio of high and low μ in data (dots) and MC (squares) as a function of E^{γ}_{T} (a), $p^{\mathrm{jet}}_{\mathrm{T,sublead}}$ (b), $\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$ (c), $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$ (d), $\beta^{\mathrm{jet}}_{\mathrm{lead}}$ (e) and β^{γ} (f) for the photon+two-jets (a,b,c,d) and the colour coherence (e,f) samples. The double ratio is shown in the lower part of the figure.*



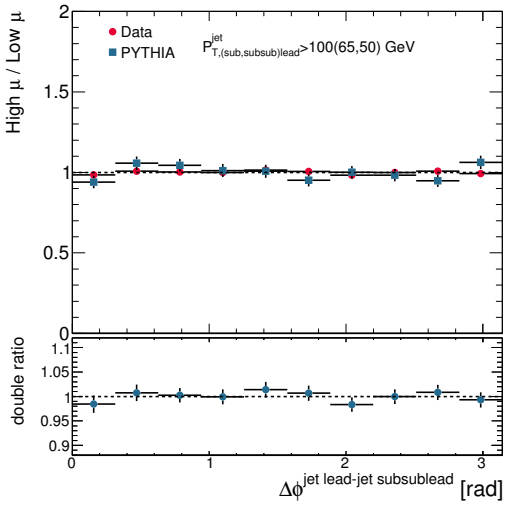
(a)



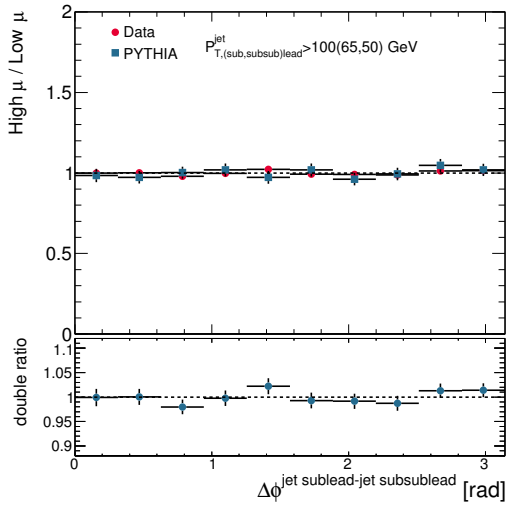
(b)



(c)



(d)



(e)

Figure 7.40: *Ratio of high and low μ in data (dots) and MC (squares) as a function of* E^{γ}_{T} *(a),* $p^{\mathrm{jet}}_{\mathrm{T,subsublead}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\ subsublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\ lead}-\mathrm{jet\ subsublead}}$ *(d),* $\Delta\phi^{\mathrm{jet\ sublead}-\mathrm{jet\ subsublead}}$ *(e) for the photon+three-jets sample. The double ratio is shown in the lower part of the figure.*

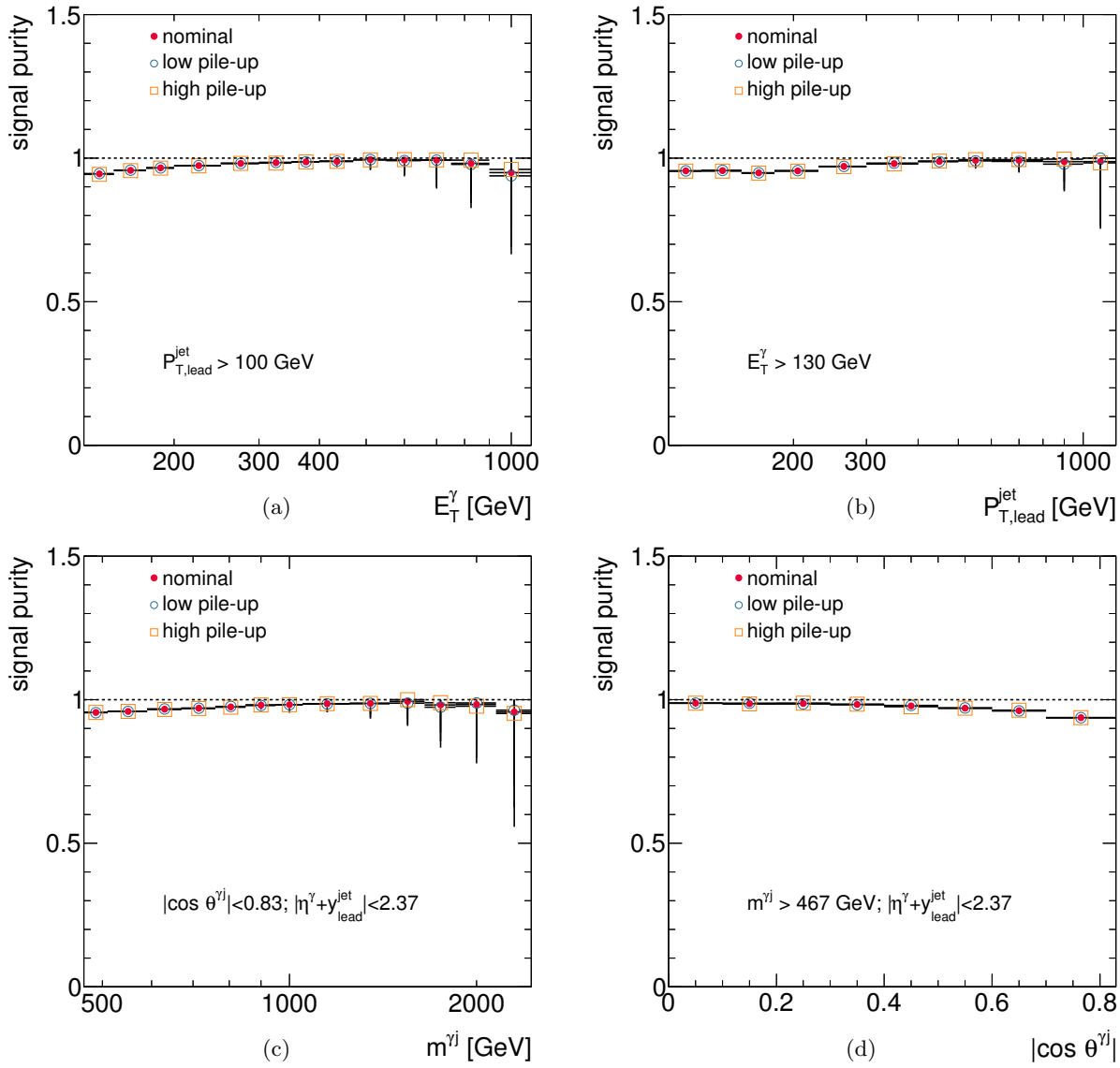


Figure 7.41: *Signal purities in data for the nominal (dots), low μ (open circles) and high μ (open squares) samples as functions of E_{T}^{γ} (a), $p_{\mathrm{T,lead}}^{\mathrm{jet}}$ (b), $m^{\gamma j}$ (c) and $|\cos\theta^{\gamma j}|$ (d) for the photon+1-jet sample.*

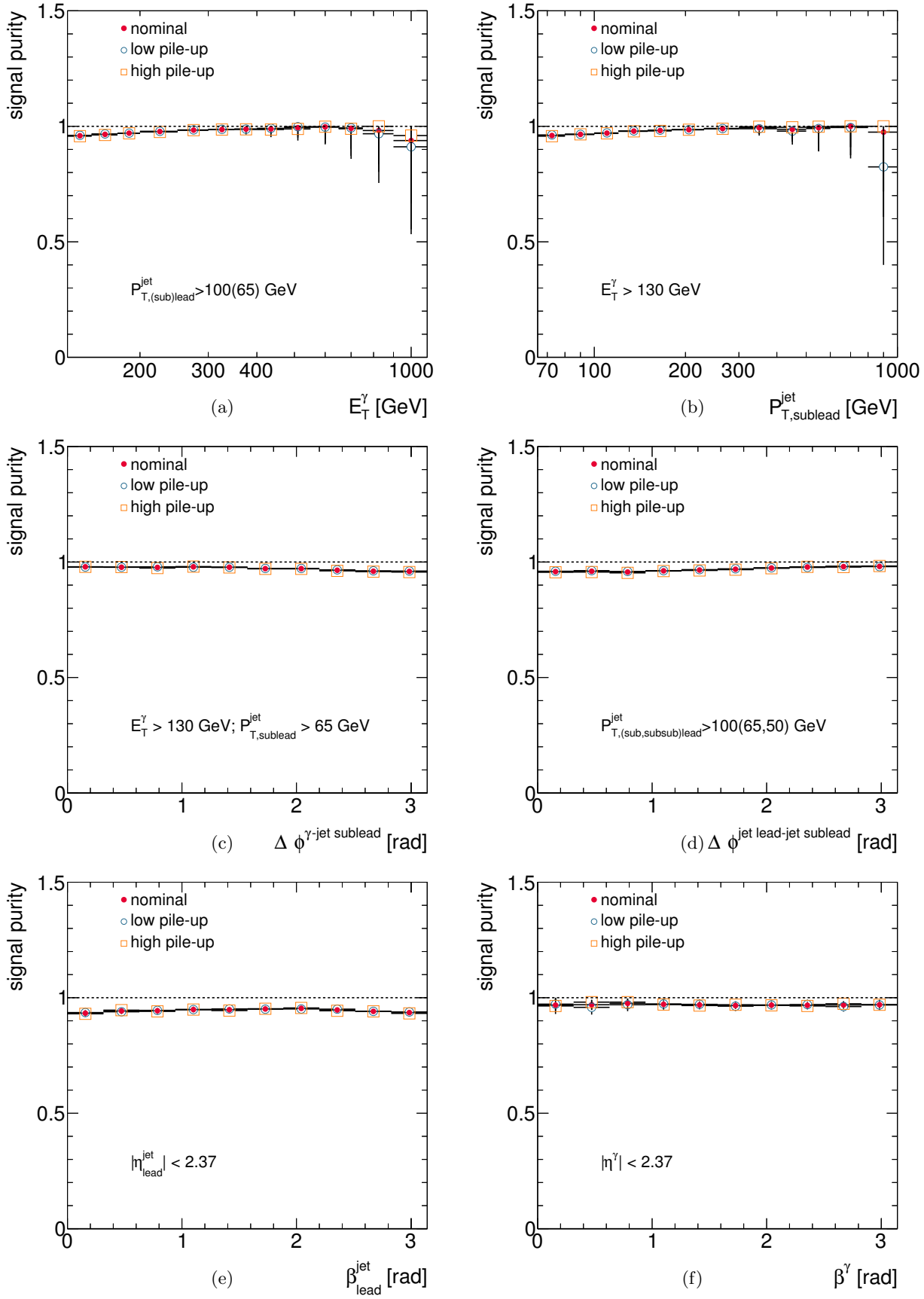


Figure 7.42: *Signal purities in data for the nominal (dots), low μ (open circles) and high μ (open squares) samples as functions of E_{T}^{γ} (a), $p_{\mathrm{T,sublead}}^{\mathrm{jet}}$ (b), $\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$ (c), $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$ (d), $\beta_{\mathrm{lead}}^{\mathrm{jet}}$ (e) and β^{γ} (f) for the photon+two-jets (a,b,c,d) and the colour coherence (e,f) samples.*

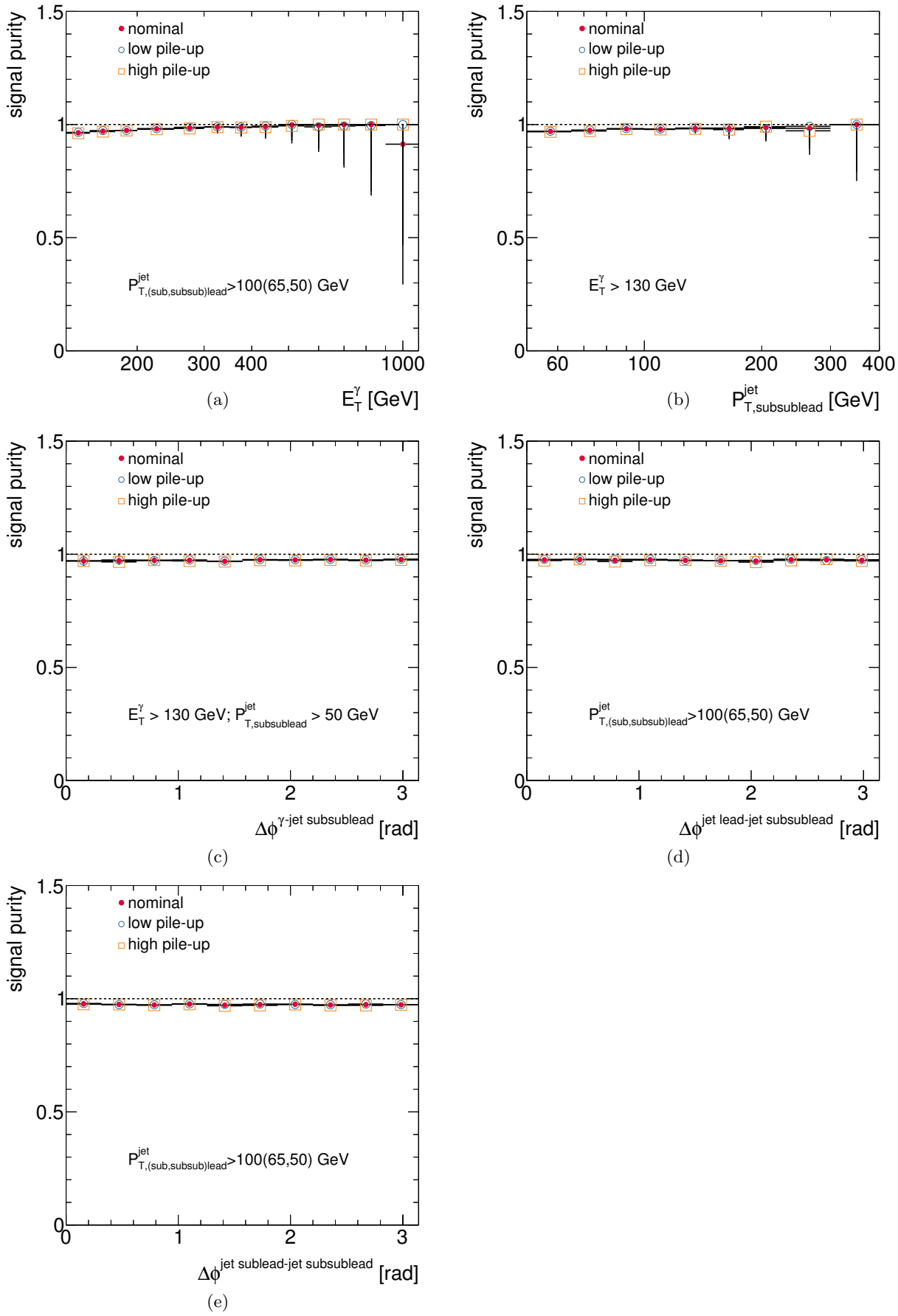


Figure 7.43: *Signal purities in data for the nominal (dots), low μ (open circles) and high μ (open squares) samples as functions of E_{T}^{γ} (a), $p_{\mathrm{T,subsublead}}^{\mathrm{jet}}$ (b), $\Delta\phi^{\gamma-\mathrm{jet\ subsublead}}$ (c), $\Delta\phi^{\mathrm{jet\ lead}-\mathrm{jet\ subsublead}}$ (d), $\Delta\phi^{\mathrm{jet\ sublead}-\mathrm{jet\ subsublead}}$ (e) for the photon+three-jets sample.*

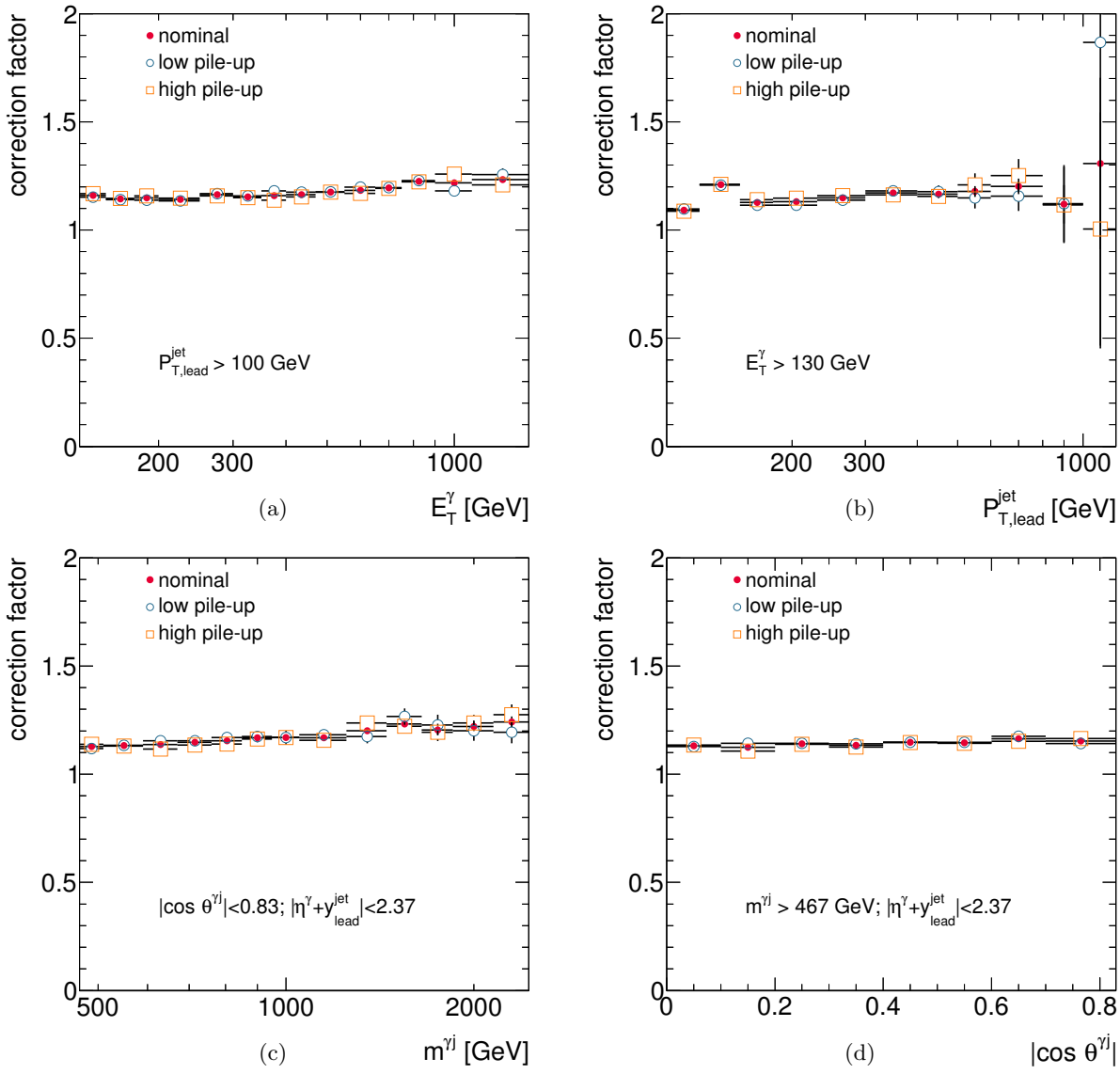


Figure 7.44: *Correction factors from* PYTHIA *for the nominal (dots), low μ (open circles) and high μ (open squares) samples as functions of E_{T}^{γ} (a), $p_{\mathrm{T,lead}}^{\mathrm{jet}}$ (b), $m^{\gamma j}$ (c) and $|\cos\theta^{\gamma j}|$ (d) for the photon+1-jet sample.*

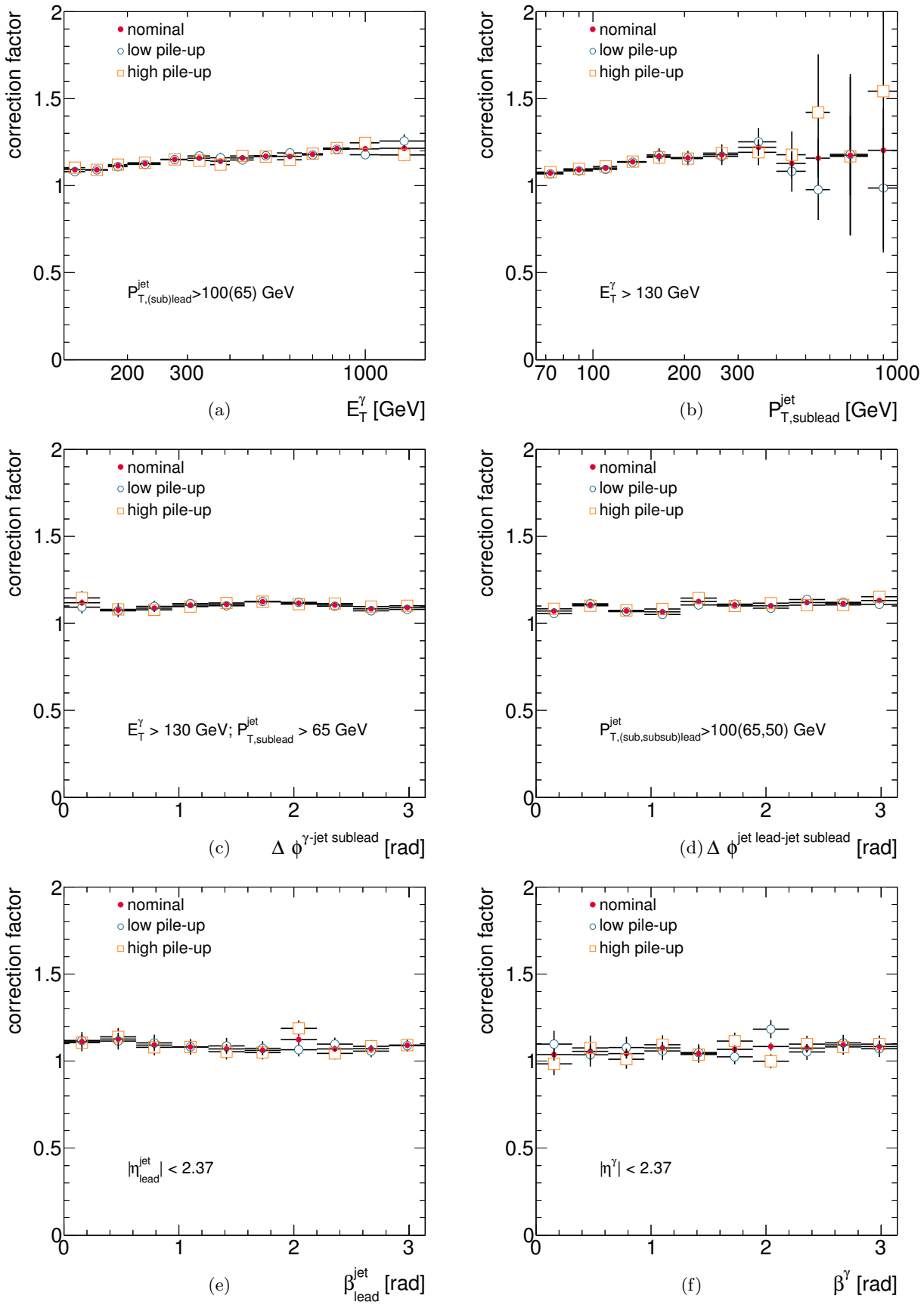


Figure 7.45: *Correction factors from* PYTHIA *(a,b,c,d) or* SHERPA *(e,f) for the nominal (dots), low μ (open circles) and high μ (open squares) samples as functions of E^{γ}_{T} (a), $p^{\mathrm{jet}}_{\mathrm{T,sublead}}$ (b), $\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$ (c), $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$ (d), $\beta^{\mathrm{jet}}_{\mathrm{lead}}$ (e) and β^{γ} (f) for the photon+two-jets (a,b,c,d) and the colour coherence (e,f) samples.*

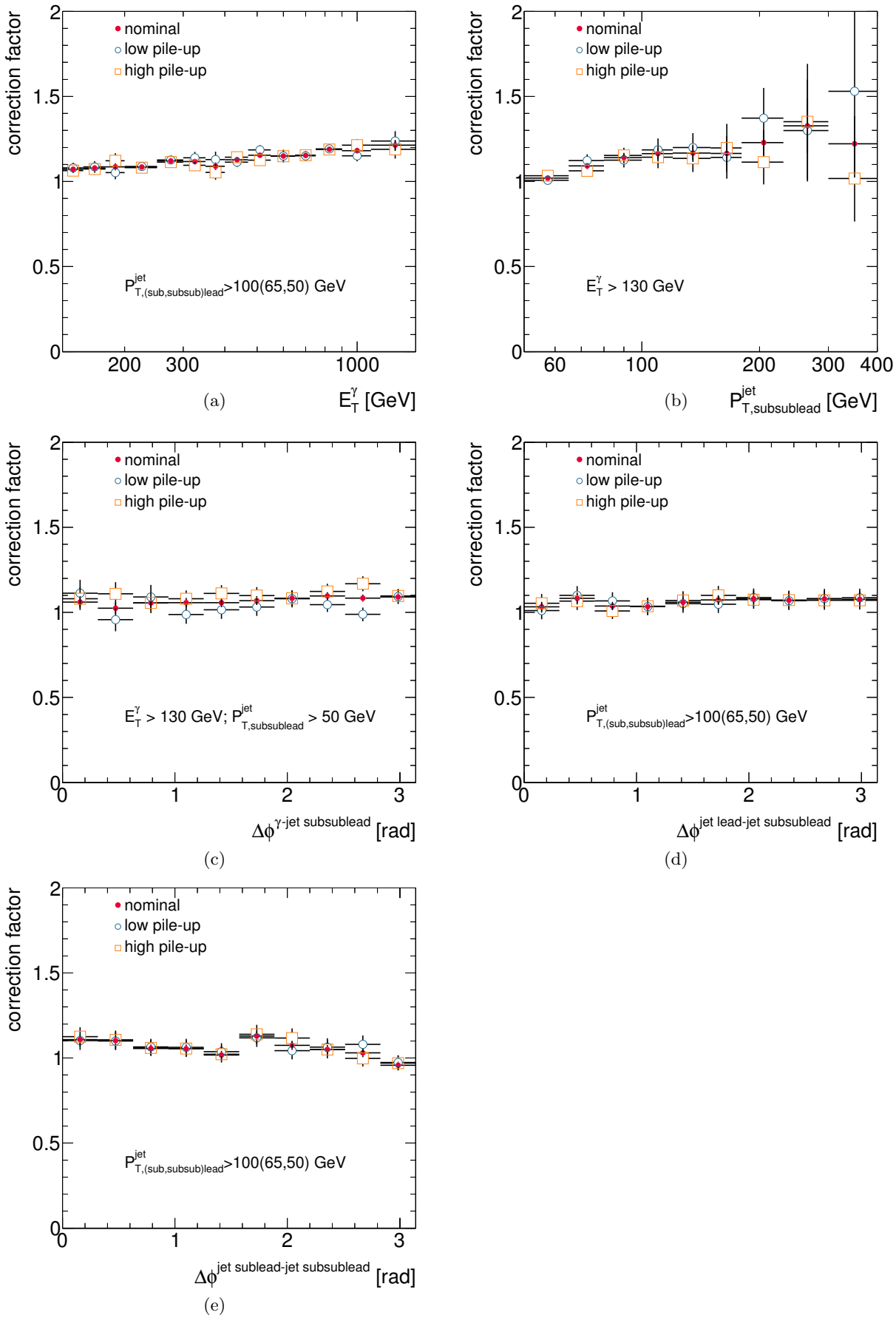


Figure 7.46: *Correction factors from* PYTHIA *for the nominal (dots), low μ (open circles) and high μ (open squares) samples as functions of E_{T}^{γ} (a), $p_{\mathrm{T,subsublead}}^{\mathrm{jet}}$ (b), $\Delta\phi^{\gamma-\mathrm{jet\ subsublead}}$ (c), $\Delta\phi^{\mathrm{jet\ lead}-\mathrm{jet\ subsublead}}$ (d), $\Delta\phi^{\mathrm{jet\ sublead}-\mathrm{jet\ subsublead}}$ (e) for the photon+three-jets sample.*

7.6 Cross checks on the additional corrections applied to the MC simulated events

Several additional corrections were applied to the MC simulations (see section 6.5). The effect on the measured cross sections was investigated by comparing the cross section before and after applying the given correction. Figures 7.47 to 7.49 show the effect of the μ reweighting, figures 7.50 to 7.52 show the effect of the $E_{\mathrm{T}}^{\mathrm{iso}}$ corrections and figures 7.53 to 7.55 show the effect of the Z-vertex reweighting on the measured cross sections. The effects on the measured cross sections are typically smaller than 1%.

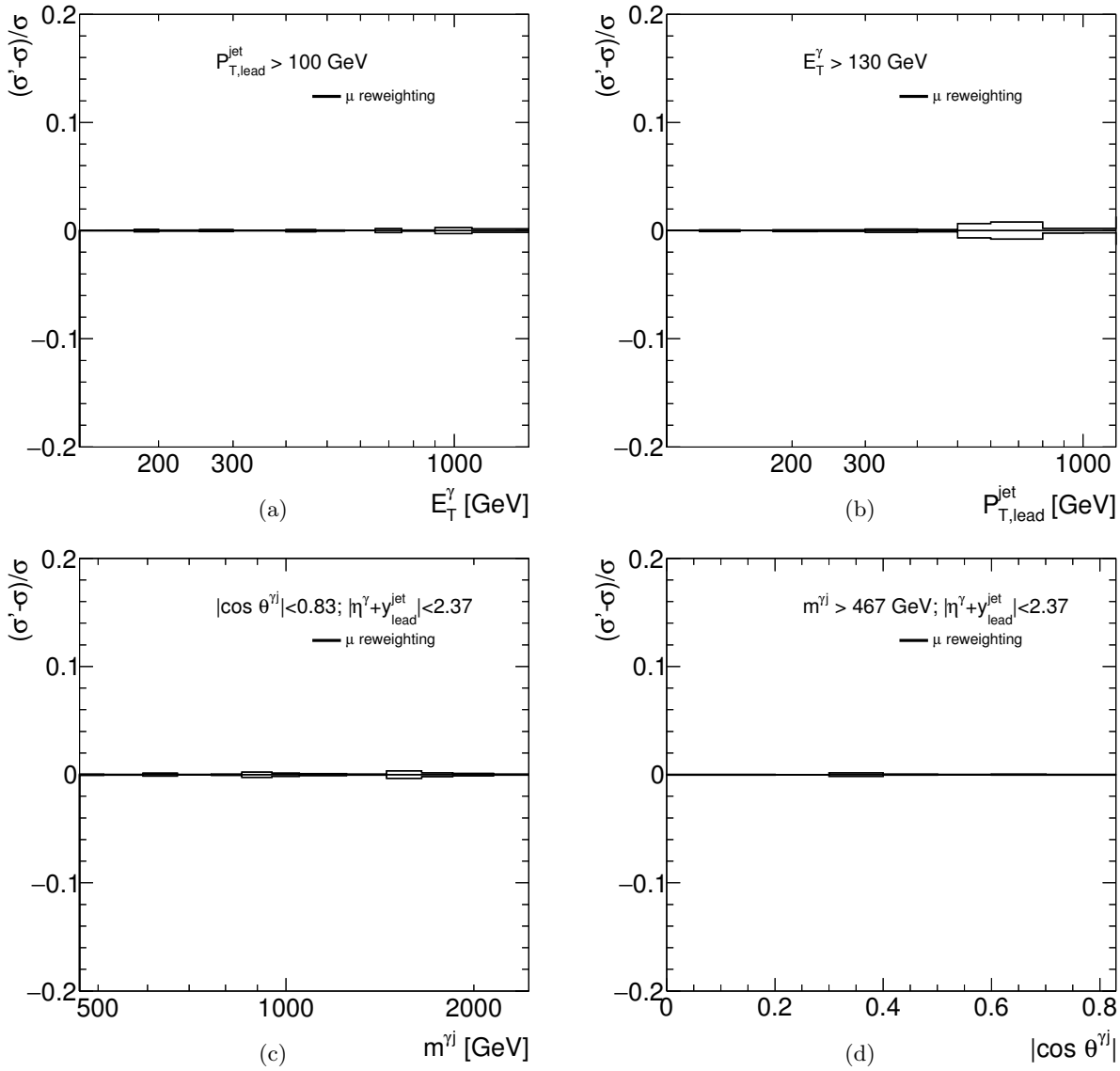


Figure 7.47: *Effect on the measured cross sections due to the μ reweighting as a function of E_{T}^{γ} (a), $p_{\mathrm{T,lead}}^{\mathrm{jet}}$ (b), $m^{\gamma j}$ (c) and $|\cos\theta^{\gamma j}|$ (d) for the photon+1-jet sample.*

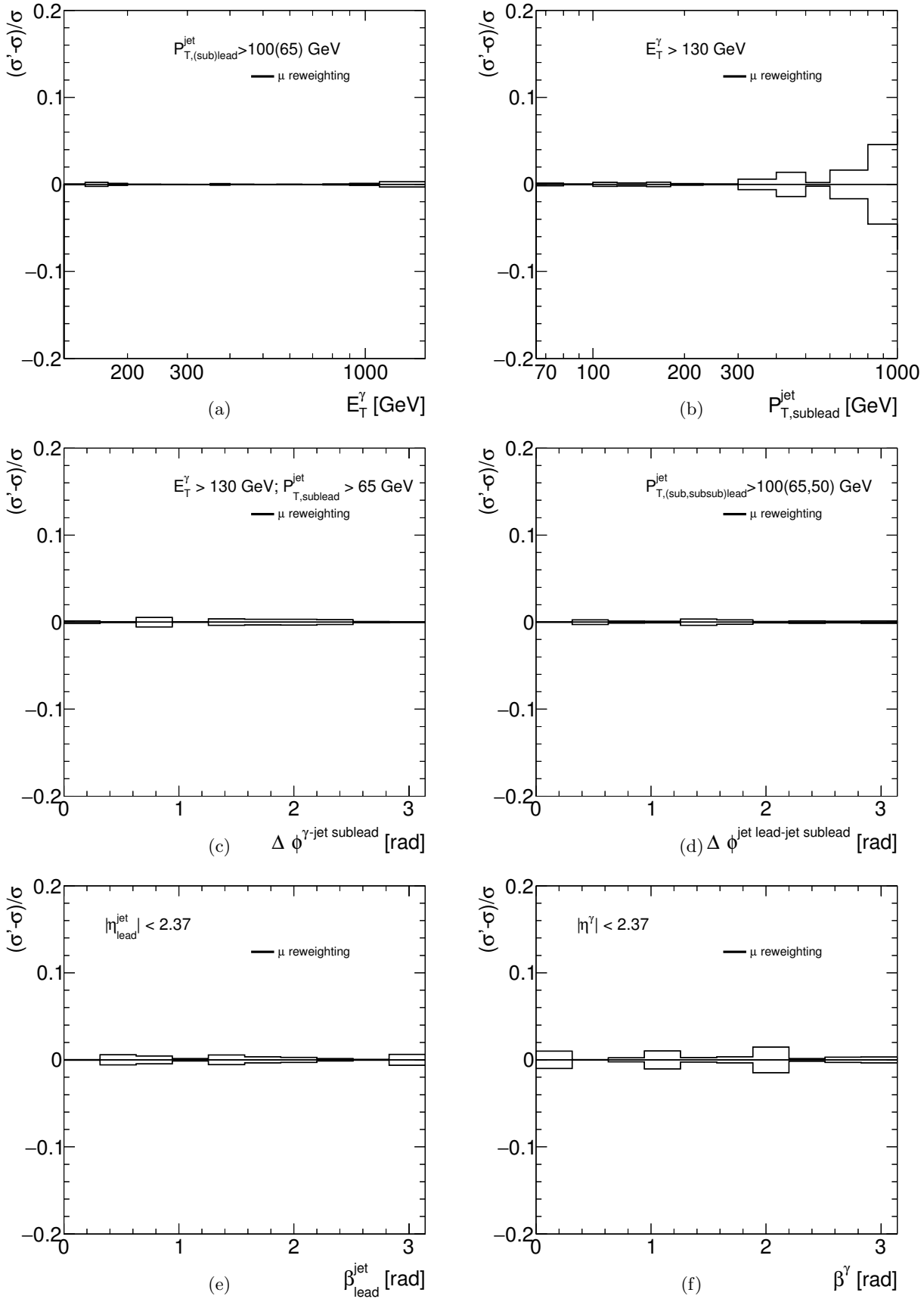


Figure 7.48: *Effect on the measured cross sections due to the μ reweighting as a function of* E^{γ}_{T} *(a),* $p^{\mathrm{jet}}_{\mathrm{T,sublead}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$ *(d),* $\beta^{\mathrm{jet}}_{\mathrm{lead}}$ *(e) and* β^{γ} *(f) for the photon+two-jets (a,b,c,d) and the colour coherence (e,f) samples.*

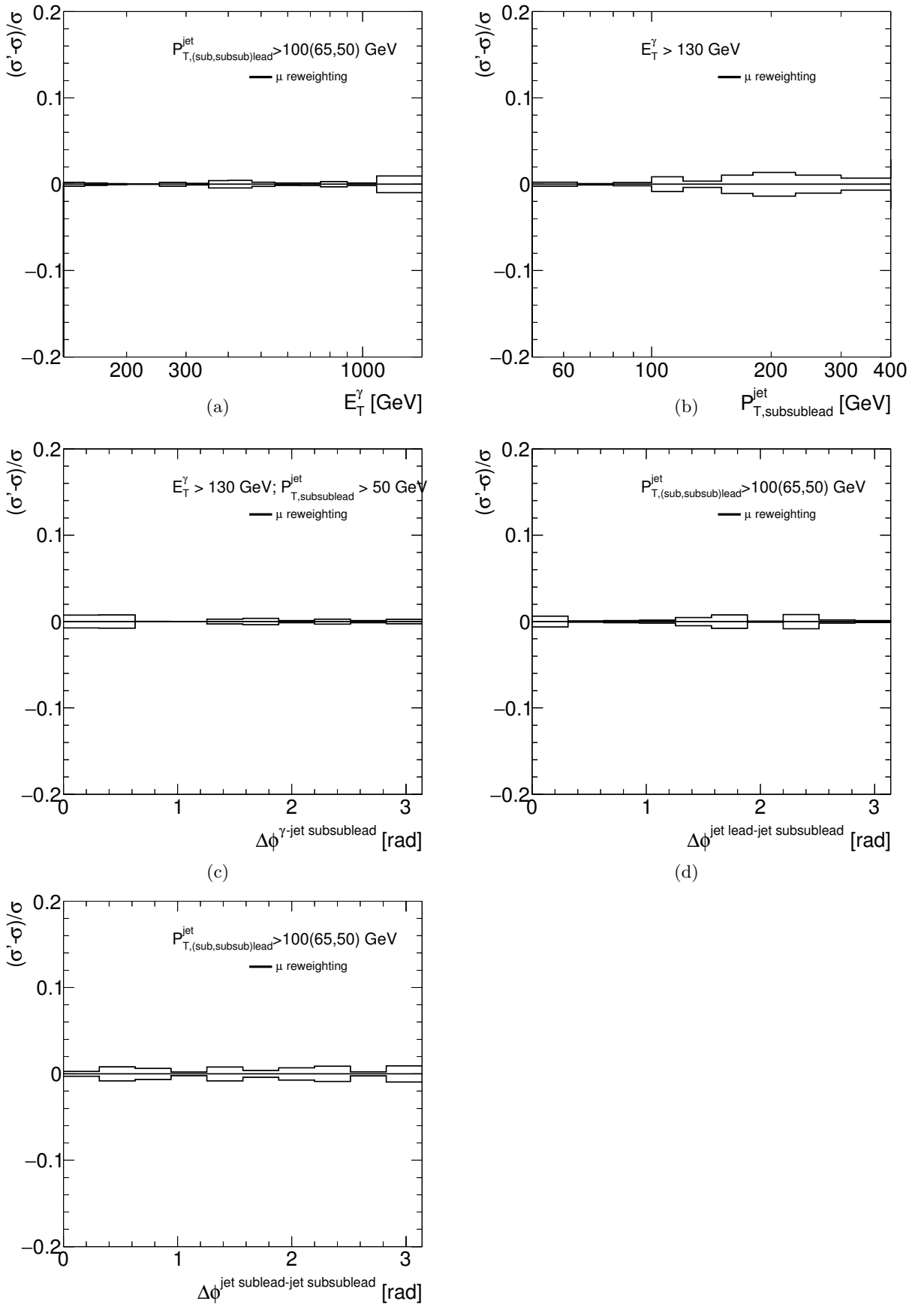


Figure 7.49: *Effect on the measured cross sections due to the μ reweighting as a function of* E^{γ}_{T} *(a),* $p^{\mathrm{jet}}_{\mathrm{T,subsublead}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\,subsublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,subsublead}}$ *(d),* $\Delta\phi^{\mathrm{jet\,sublead}-\mathrm{jet\,subsublead}}$ *(e) for the photon+three-jets sample.*

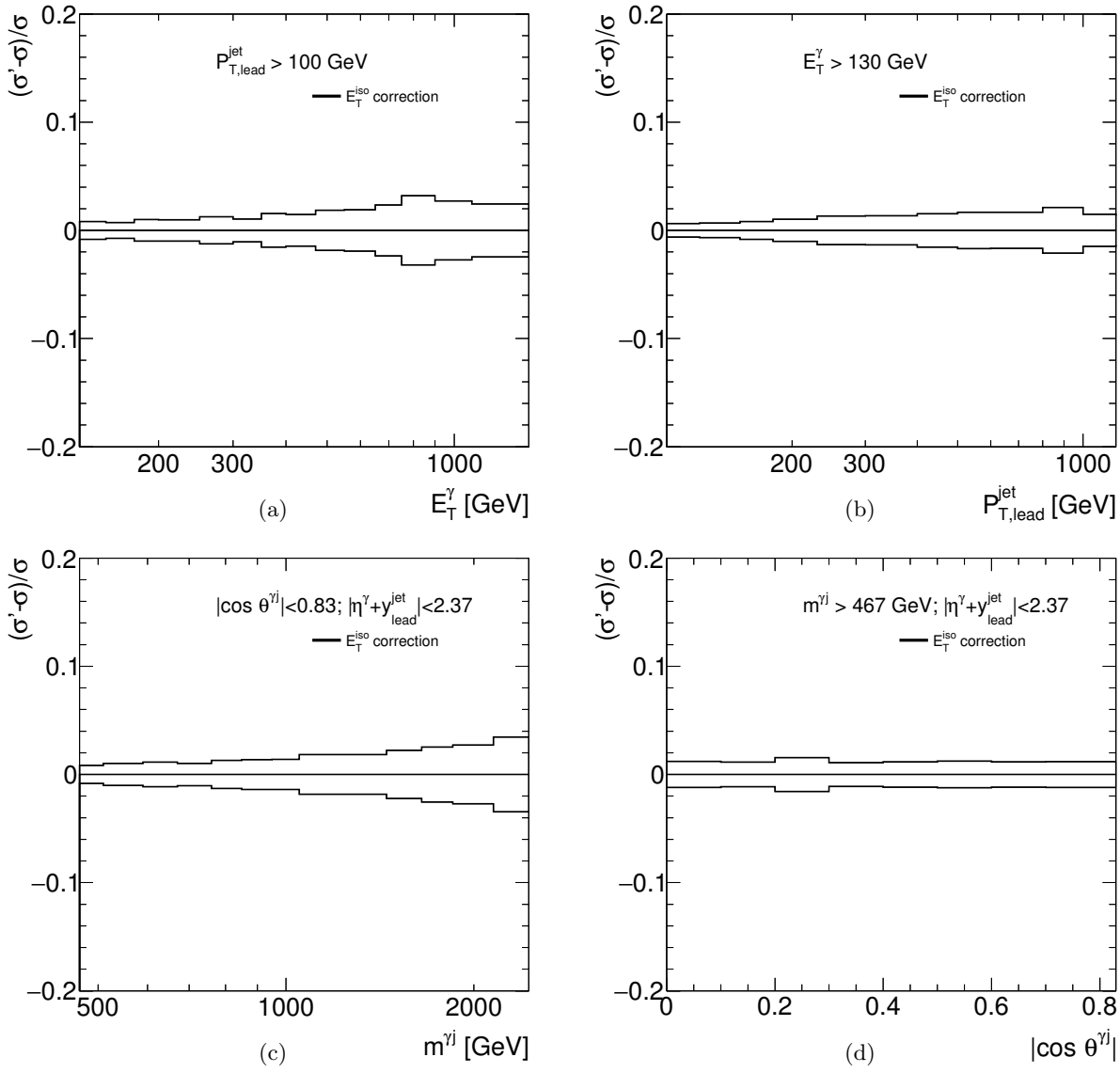


Figure 7.50: *Effect on the measured cross sections due to the $E_{\mathrm{T}}^{\mathrm{iso}}$ corrections as a function of E_{T}^{γ} (a), $p_{\mathrm{T,lead}}^{\mathrm{jet}}$ (b), $m^{\gamma j}$ (c) and $|\cos\theta^{\gamma j}|$ (d) for the photon+1-jet sample.*

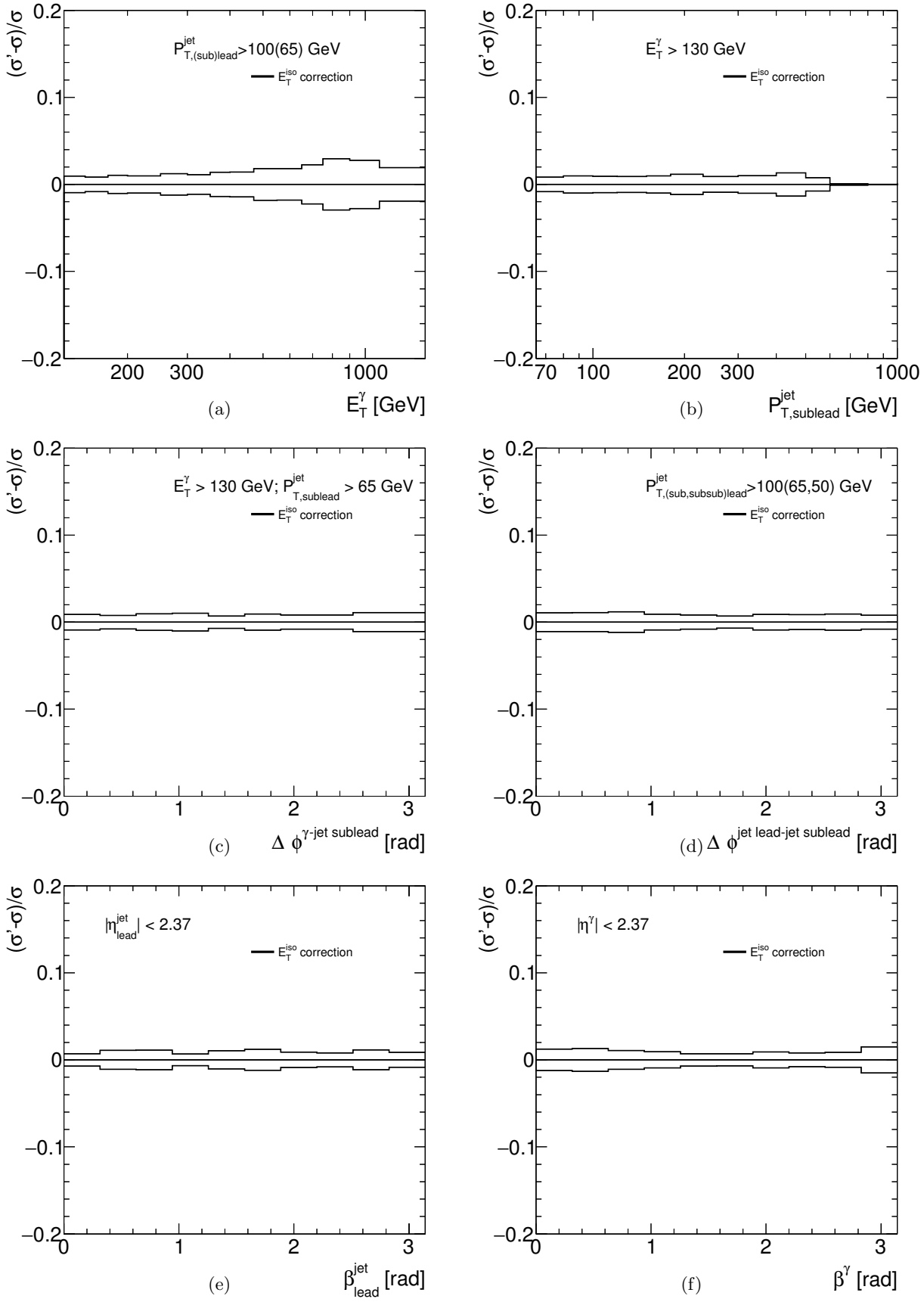


Figure 7.51: *Effect on the measured cross sections due to the $E_{\mathrm{T}}^{\mathrm{iso}}$ corrections as a function of E_{T}^{γ} (a), $p_{\mathrm{T,sublead}}^{\mathrm{jet}}$ (b), $\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$ (c), $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$ (d), $\beta_{\mathrm{lead}}^{\mathrm{jet}}$ (e) and β^{γ} (f) for the photon+two-jets (a,b,c,d) and the colour coherence (e,f) samples.*

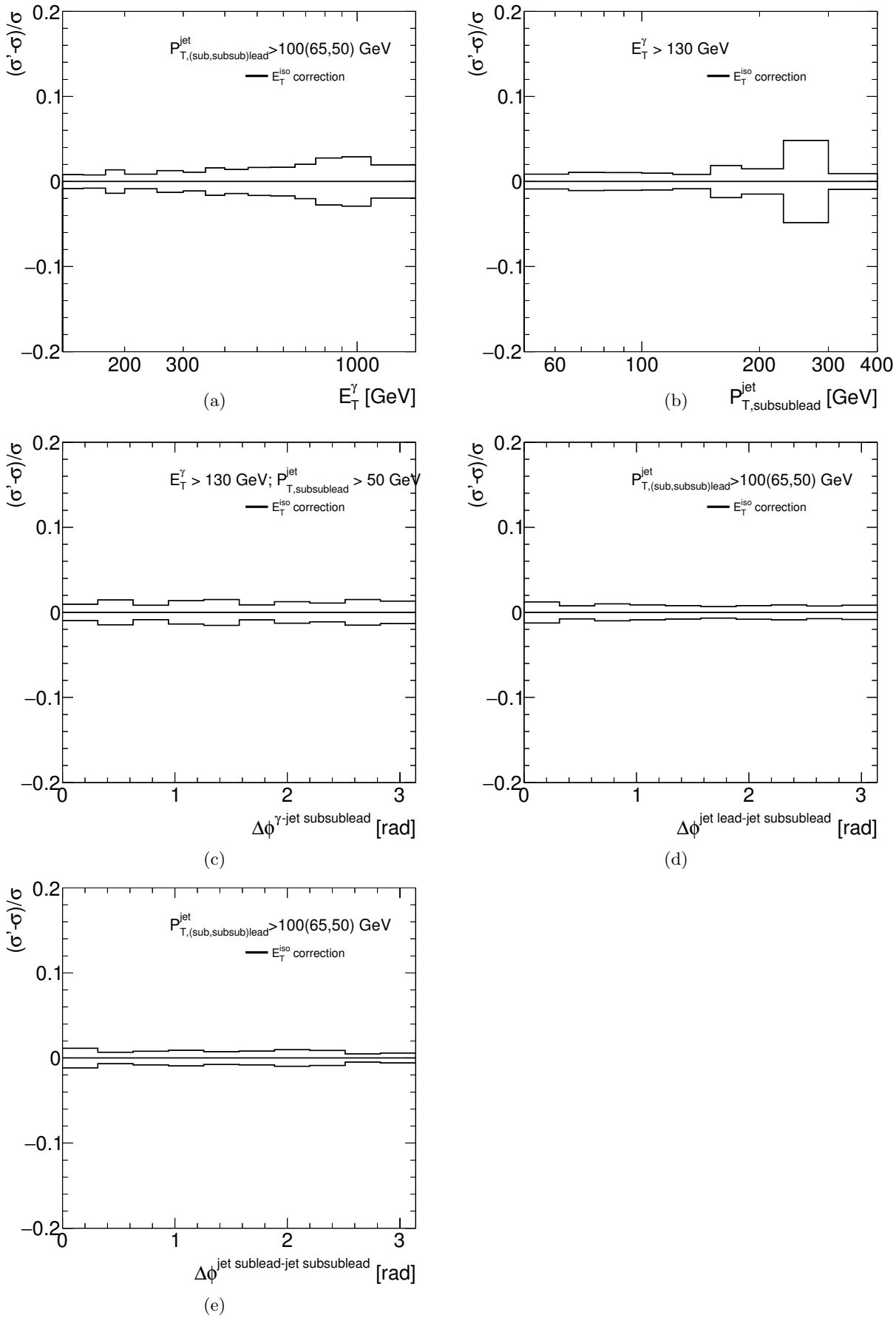


Figure 7.52: *Effect on the measured cross sections due to the* $E_{\mathrm{T}}^{\mathrm{iso}}$ *corrections as a function of* E_{T}^{γ} *(a),* $p_{\mathrm{T,subsublead}}^{\mathrm{jet}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\,subsublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,subsublead}}$ *(d),* $\Delta\phi^{\mathrm{jet\,sublead}-\mathrm{jet\,subsublead}}$ *(e) for the photon+three-jets sample.*

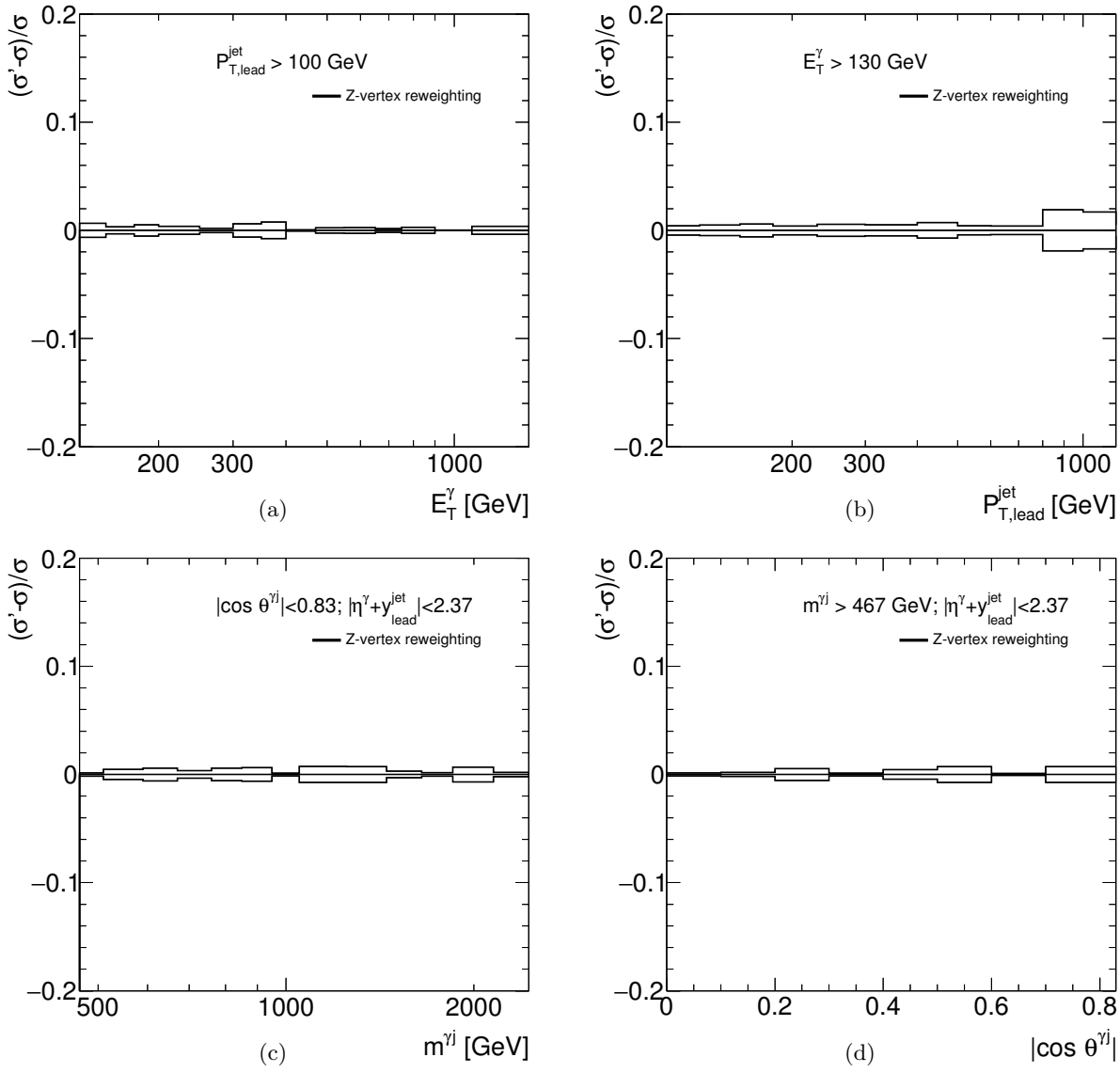


Figure 7.53: *Effect on the measured cross sections due to the Z-vertex reweighting as a function of* E_{T}^{γ} *(a),* $p_{\mathrm{T,lead}}^{\mathrm{jet}}$ *(b),* $m^{\gamma j}$ *(c) and* $|\cos\theta^{\gamma j}|$ *(d) for the photon+1-jet sample.*

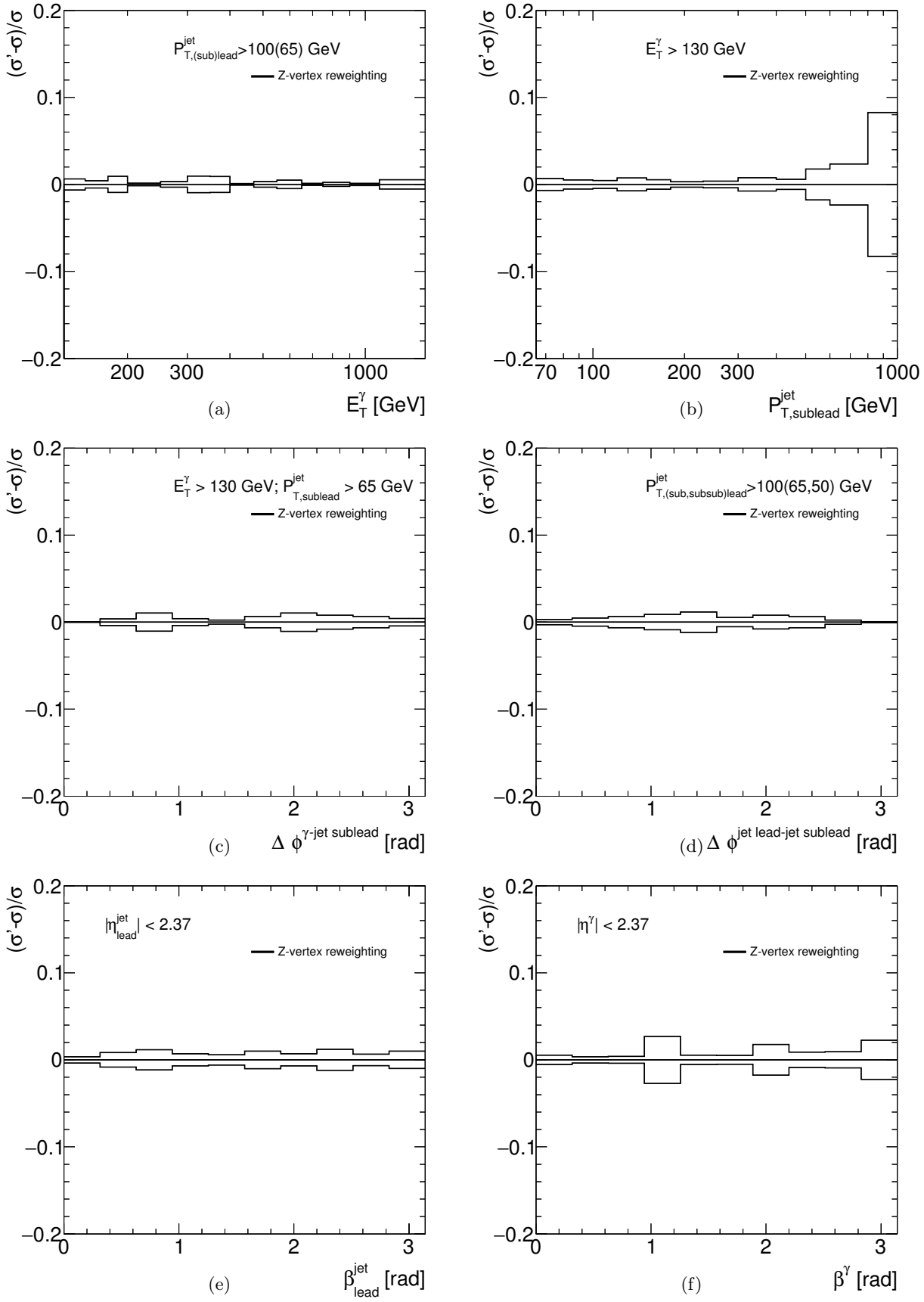


Figure 7.54: *Effect on the measured cross sections due to the Z-vertex reweighting as a function of* E^{γ}_{T} *(a),* $p^{\mathrm{jet}}_{\mathrm{T,sublead}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$ *(d),* $\beta^{\mathrm{jet}}_{\mathrm{lead}}$ *(e) and* β^{γ} *(f) for the photon+two-jets (a,b,c,d) and the colour coherence (e,f) samples.*

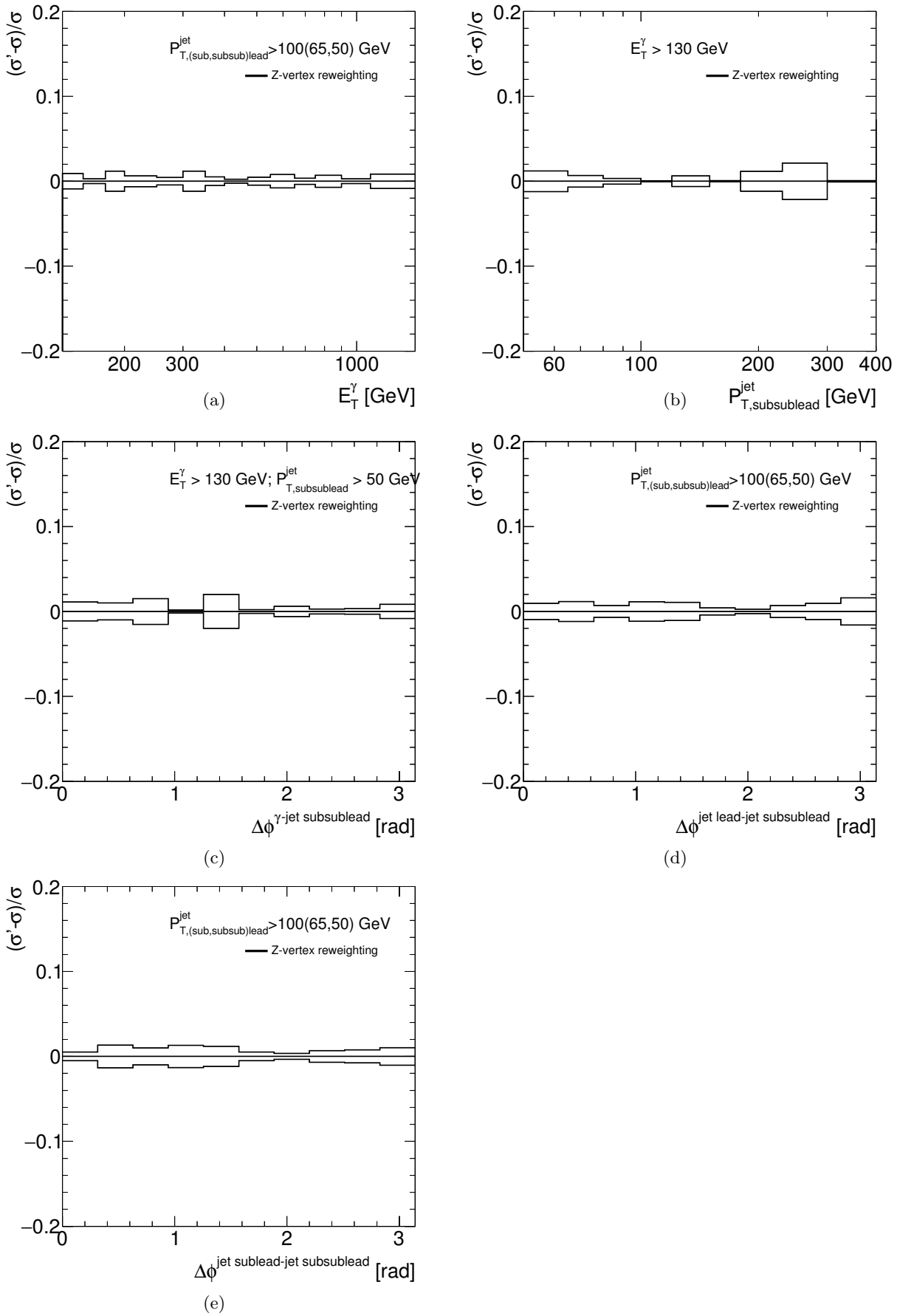


Figure 7.55: *Effect on the measured cross sections due to the Z-vertex reweighting as a function of* E^{γ}_{T} *(a),* $p^{\mathrm{jet}}_{\mathrm{T,subsublead}}$ *(b),* $\Delta\phi^{\gamma-\mathrm{jet\ subsublead}}$ *(c),* $\Delta\phi^{\mathrm{jet\ lead}-\mathrm{jet\ subsublead}}$ *(d),* $\Delta\phi^{\mathrm{jet\ sublead}-\mathrm{jet\ subsublead}}$ *(e) for the photon+three-jets sample.*

Summary and conclusions

In this dissertation, two analyses, involving photons, jets and top quarks and performed using data collected by the ATLAS detector at the LHC, have been described. These analyses provided stringent tests of the Standard Model (SM) using isolated photons in association with jets and a novel method to identify boosted particles.

A new method to identify highly boosted massive particles based on their hadronic decay mode has been presented. This method is based on the reconstruction of jets with distance parameter $R = 1.5$ using the anti-k_{T} algorithm ("large-R jets"), the application of the k_{T} algorithm on the constituents of these large-R jets to reconstruct subjets and the subsequent application of the anti-k_{T} algorithm with distance parameter $R = 0.4$ on the constituents of each of these subjets to obtain the so called small-R jets. This last step removes successfully and in an infrared- and collinear-safe way the soft contributions from the underlying event and pileup, inherent to pp collisions. The small-R jets characterise the hard-constituents inside the large-R jets. The mehod relies only on the knowledge of the number of hadronic decay products of the boosted particle and thus it can be easily used in searches of SM or BSM hadronically decaying massive particles.

The method has been tested in the identification of boosted top-quark pair ($t\bar{t}$) production in the semi-leptonic channel at a centre-of-mass energy of 7 TeV, using the data collected by ATLAS in 2011 and corresponding to an integrated luminosity of 4.6 fb^{-1}. The events were selected by requiring the presence of an electron with transverse energy (E_{T}) higher than 40 GeV and a large-R jet with transverse momentum p_{T} higher than 400 GeV. Monte Carlo (MC) samples for the signal and the most relevant backgrounds were generated using various MC programs to study the properties of the signal and background processes, respectively. A method to obtain the optimal admixture of the different processes has been developed, based on a multi-variable simultaneous fits to the data. This optimal admixture was used to estimate the background to $t\bar{t}$ production.

Three small-R jets, corresponding to the three partons of the hadronically decaying top-quark, were obtained from the constituents of the large-R and the signal to background ratio was enhanced by exploiting the differences in the topology of the small-R jets between the signal and the background. In addition, the hadronically decaying top quark was reconstructed using the small-R jets and the resulting distributions as functions of the invariant masses of three (M_{JJJ}) and two (M_{JJ}) small-R jets for the selected data events show peaks around the expected values of M_{top} and M_W, respectively, with a very good resolution.

After background subtraction, the data distributions are well described by the $t\bar{t}$ MC samples, both in shape and normalisation. In addition, a high-purity sample, based on a b-tagging technique, was used to cross-check the identification method. The distributions as functions of M_{JJJ} and M_{JJ}, after background subtraction, with and without b-tagging have very similar shapes; this supports the fact that the method developed identifies properly events originating from $t\bar{t}$ production.

The efficiency for identifying the boosted top quarks was estimated for this novel method using a sample of signal MC events and found to be 77%. The rejection factor for several sources

of background was also obtained from background MC samples and found to be 5.9 for W+jets, 5.3 for Z+jets, 3.1 for single top and 6.7 for QCD processes. The efficiency and rejection factors obtained are similar or superior to the ones obtained by several top-taggers already in use by the ATLAS Collaboration. In addition, the global efficiency of the analysis in the identification of $t\bar{t}$ events in the semi-leptonic channel was found to be 60%.

Tests of the dynamics and colour-coherence effects in isolated photons in association with jets have been presented. Measurements of differential cross sections at a centre-of-mass energy of 8 TeV using the data collected by ATLAS, corresponding to an integrated luminosity of 20.3 fb^{-1}, have been performed.

A sample of isolated photon plus one jet events has been selected by requiring the presence of an isolated photon with transverse energy (E_{T}^{γ}) higher than 130 GeV in addition to at least one jet with $p_{\mathrm{T}} > 100$ GeV. The isolation was ensured by limiting the amount of energy in a cone of $R = 0.4$ around the photon ($E_{\mathrm{T}}^{\mathrm{iso}}$) to be below 10 GeV. The jets were reconstructed using the anti-k_{T} jet algorithm with distance parameter $R = 0.6$. Sub-samples of events with two and three jets have also been selected by requiring the presence of additional jets with $p_{\mathrm{T}} > 65$ GeV and $p_{\mathrm{T}} > 50$ GeV, respectively.

Monte Carlo samples of prompt photon plus jets production have been generated with two different MC programs, Pythia and Sherpa. The MC samples were used to calculate the correction factors used in the bin-by-bin unfolding, to evaluate certain systematic uncertainties and to estimate the hadronisation and underlying-event effect corrections to the next-to-leading-order (NLO) calculations.

The background was subtracted using a data-driven technique, based on a "2D-sideband" method, which relies on signal-depleted control regions defined on a plane of two variables which are uncorrelated for the background. This plane was defined using a photon identification variable (which contains information related to the shower-shape of the photon) and the $E_{\mathrm{T}}^{\mathrm{iso}}$ variable. The method takes also into account the expected number of signal events in the control regions, which is estimated from the MC samples. The purity as a function of all the measured observables is always above 90%.

Differential cross sections were measured for the photon+one-jet sample as functions of E_{T}^{γ}, the transverse momentum of the leading jet ($p_{\mathrm{T,lead}}^{\mathrm{jet}}$), the invariant mass of the photon-jet system ($m^{\gamma j}$) and $|\cos\theta^{\gamma j}|$, which corresponds to the cosine of the scattering angle in the centre-of-mass frame of the photon and the jet. The measured cross sections have been compared to NLO calculations obtained using Jetphox, corrected for hadronisation and underlying-event effects. The NLO predictions provide a very good description of the data. The most relevant theoretical uncertainties were evaluated, such as the effect of terms beyond NLO, the effect of the uncertainty in the parametrisation of the proton PDFs and the uncertainty on the value of α_s.

The shape of the measured differential cross section as a function of $|\cos\theta^{\gamma j}|$ is consistent with that of a process dominated by quark-exchange. To gain further insight into the dynamics of the photon-jet system, the scale evolution of the differential cross section as a function of $|\cos\theta^{\gamma j}|$ was measured in different regions of $m^{\gamma j}$. The measurements show an increase of the value for $|\cos\theta^{\gamma j}| = 0.8$ as $m^{\gamma j}$ increases, which is well described by the NLO calculations.

Differential cross sections were measured for the photon+two-jet and photon+three-jet samples as functions of E_{T}^{γ}, the transverse momentum of the sub-leading jet ($p_{\mathrm{T,sublead}}^{\mathrm{jet}}$), the transverse momentum of the sub-sub-leading jet ($p_{\mathrm{T,subsublead}}^{\mathrm{jet}}$) and the differences in the azimuthal angle between the photon and the jets and the jets between themselves ($\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$, $\Delta\phi^{\gamma-\mathrm{jet\,subsublead}}$, $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$, $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,subsublead}}$ and $\Delta\phi^{\mathrm{jet\,sublead}-\mathrm{jet\,subsublead}}$). The measured cross sections were compared to the predictions of Sherpa and Pythia. The predictions of Sherpa provide a good description of the data, in particular, the differences in the azimuthal angles

in the photon+two-jet samples, while PYTHIA provides a better description of the differential cross section as a function of E_{T}^{γ} for $E_{\mathrm{T}}^{\gamma} < 900$ GeV. The poorer performance of PYTHIA when dealing with additional jets is explained by the fact that in the $2 \to 2$ predictions of PYTHIA, a second jet can arise only from the parton shower, whereas in SHERPA, $2 \to 3$ and $2 \to 4$ matrix elements contributions can provide a second hard jet.

The scale evolution for the photon+two-jets and photon+three-jets samples has been studied by measuring the differential cross sections as functions of the differences in the azimuthal angle between the objects in two different regions of E_{T}^{γ} ($130 < E_{\mathrm{T}}^{\gamma} < 300$ GeV and $E_{\mathrm{T}}^{\gamma} > 300$ GeV). In all cases, the lower E_{T}^{γ} region behaves differently than the full region: for $\Delta\phi^{\gamma-\text{jet sublead}}$ and $\Delta\phi^{\gamma-\text{jet subsublead}}$ the cross sections are enhanced for large $\Delta\phi$ while $\Delta\phi^{\text{jet lead}-\text{jet sublead}}$, $\Delta\phi^{\text{jet lead}-\text{jet subsublead}}$ and $\Delta\phi^{\text{jet sublead}-\text{jet subsublead}}$ are suppressed.

Colour-coherence effects, in which colour connection between initial and/or final-state partons modify the pattern of radiation around coloured partons, were measured for the first time by ATLAS in photon+jets events. Two sub-samples of events with at least two jets in which the sub-leading jet is either close to the photon or the leading-jet were defined with that purpose. Differential cross sections were measured using such samples as functions of the angle in the $\eta-\phi$ plane between the sub-leading jet and the close-by photon (β^{γ}) or between the sub-leading jet and the close-by leading jet ($\beta_{\mathrm{lead}}^{\mathrm{jet}}$). Such cross sections provide a measurement of the direction of the radiation around both objects. Both cross sections display a very different behaviour: the measured $d\sigma/d\beta_{\mathrm{lead}}^{\mathrm{jet}}$ increases as $\beta_{\mathrm{lead}}^{\mathrm{jet}}$ increases, slower for $\beta_{\mathrm{lead}}^{\mathrm{jet}} < \pi/2$ and faster for $\beta_{\mathrm{lead}}^{\mathrm{jet}} > \pi/2$, whereas $d\sigma/d\beta^{\gamma}$ exhibits a maximum at $\beta^{\gamma} \approx \pi/2$. The predictions of SHERPA provide a good description of the measured cross sections. To ascertain the presence of colour-coherence effects from data alone, the ratio of these measured cross sections was performed; enhancements of the cross section as a function of $\beta_{\mathrm{lead}}^{\mathrm{jet}}$ at $\beta_{\mathrm{lead}}^{\mathrm{jet}} = 0$ and at $\beta_{\mathrm{lead}}^{\mathrm{jet}} = \pi$ with respect to the cross section as a function of β^{γ} are clearly observed. This behaviour indicates an enhancement of radiation in the direction of the initial-state partons around the leading jet with respect to a non-coloured particle, which is consistent with the effect of colour coherence.

In summary, two analyses have been presented which improve our knowledge of the fundamental particles and interactions at the energy frontier provided by the LHC. The photon plus jets cross-section measurements and their comparison to theoretical predictions provide a stringent test of QCD and yield further information to characterise one of the most important backgrounds to Higgs decaying into two photons and any other new particle decaying into photons. The new method developed to identify boosted particles in an infrared- and collinear-safe way provides a reliable approach for searches of new very massive particles.

Resumen en castellano

En esta tesis se han descrito dos análisis sobre fotones, jets y quarks top, realizados usando datos obtenidos con el detector ATLAS. Dichos análisis han proporcionado tests rigurosos del modelo estándar usando fotones aislados producidos en asociación con jets y un novedoso método para identificar partículas con un alto momento transverso.

Se ha presentado un nuevo método para identificar partículas con un alto momento transverso basado en su modo de desintegración hadrónico. Este método esta basado en la reconstrucción de jets con un radio de 1.5 usando el algoritmo anti-k_{T} (jets de radio grande), la utilización del algoritmo k_{T} sobre los jets de radio grande para reconstruir subjets y la subsiguiente utilización del algoritmo anti-k_{T} con un radio de 0.4 sobre los constituyentes de cada uno de dichos subjets para obtener lo que se denominan jets de radio pequeño. Este último paso elimina las contribuciones suaves debidas al suceso subyacente y a otras colisiones simultáneas, siempre presentes en las colisiones protón-proton, de una manera que no presenta singularidades en los límites infrarrojo o colineal. Los jets de radio pequeño así obtenidos caracterizan los productos de desintegración dentro del jet de radio grande. Este método parte tan sólo de asumir un número determinado de partones provenientes de la partícula con un alto momento transverso y por lo tanto se puede usar para cualquier búsqueda de partículas masivas que decaigan hadrónicamente en el modelo estándar o en física más allá del modelo estándar.

Dicho método ha sido investigado mediante la identificación de parejas de quarks top ($t\bar{t}$) con un alto momento transverso en el canal de desintegración semi-leptónico, usando para ello los datos obtenidos con el detector ATLAS durante 2011 a una energía del centro de masas de 7 TeV. Dicho conjunto de datos corresponde a una luminosidad integrada de 4.6 fb^{-1}. Los sucesos se han seleccionado requiriendo la presencia de un electrón con una energía transversa de al menos 40 GeV y la de un jet de radio grande con un momento transverso de al menos 400 GeV. Muestras de Monte Carlo (MC) para la señal y las fuentes de ruido de fondo más relevantes han sido generadas y utilizadas para estudiar las propiedades de los procesos implicados. Utilizando dichas muestras se ha desarrollado un método para obtener la mezcla óptima de los diferentes procesos, basada en un ajuste multivariable simultáneo a las distribuciones de datos. Esta mezcla óptima se ha utilizado para estimar las fuentes de ruido de fondo a la producción de $t\bar{t}$.

Utilizando el método descrito en esta tesis, se han identificado tres jets de radio pequeño, correspondientes a los tres partones provenientes del top quark que se desintegra hadrónicamente, utilizando los constituyentes del jet de radio grande y se ha mejorado el ratio entre la señal y el ruido de fondo aprovechando las diferencias en la topología de dichos jets entre ambos tipos de procesos. Además, se ha reconstruido el top quark que se desintegra hadrónicamente utilizando esos mismos jets; las distribuciones resultantes como función de la masa invariante de tres (M_{JJJ}) y dos (M_{JJ}) jets de radio pequeño para los sucesos seleccionados presentan picos alrededor de los valores de la masa del quark top (M_{top}) y de la masa del bosón W (M_W) respectivamente, con una muy buena resolución.

Una vez substraído el ruido de fondo restante, las distribuciones de datos están bien descritas por las muestras MC de $t\bar{t}$, tanto en forma como en normalización. Además, se ha seleccionado una muestra de alta pureza, basándose en una técnica de identificación de quarks b, para realizar

una comprobación del método de identificación. Las distribuciones de datos como función de M_{JJJ} y M_{JJ}, después de sustraer el ruido de fondo, con y sin la identificación de quarks b adicional, tienen una forma muy similar; esto demuestra que el método desarrollado es capaz de identificar de forma correcta sucesos provenientes de la producción de $t\bar{t}$.

Se ha estimado la eficiencia para la identificación de top quarks de alto momento transverso utilizando una muestra de MC de la señal, siendo ésta de 77%. Se ha obtenido también el factor de supresión para diversas fuentes de ruido de fondo, utilizando muestras de MC de los distintos procesos, siendo éste 5.9 para producción de W+jets, 5.3 para Z+jets, 3.1 para producción de quark top en solitario y 6.7 para procesos de QCD. La eficiencia y los factores de supresión obtenidos son similares o mejores que los obtenidos con otras técnicas de identificación de quarks top que se usan habitualmente en la colaboración ATLAS. Finalmente, se ha obtenido la eficiencia del análisis presentado en la identificación de sucesos de producción de $t\bar{t}$ en el canal de desintegración semi-leptónico, siendo ésta 60%.

Se han presentado tests de la dinámica y de la coherencia de color en fotones aislados producidos en asociación con jets mediante medidas de secciones eficaces diferenciales, usando para ello los datos obtenidos con el detector ATLAS durante 2012 a una energía en el centro de masas de 8 TeV. Dicho conjunto de datos se corresponde con una luminosidad integrada de 20.3 fb^{-1}.

Se ha seleccionado una muestra de fotones aislados en asociación con un jet requiriendo la presencia de un fotón aislado con una energía transversa (E_{T}^{γ}) superior a 130 GeV y la de al menos un jet con un momento transverso de al menos 100 GeV. Se ha impuesto el aislamiento limitando la energía transversa en un cono de radio 0.4 alrededor del fotón ($E_{\mathrm{T}}^{\mathrm{iso}}$) a un máximo de 10 GeV. Los jets han sido reconstruidos usando el algoritmo anti-k_{T} con un radio de 0.6. Se han seleccionado también sub-muestras de sucesos con dos y tres jets requiriendo la presencia de jets adicionales con un momento transverso superior a 65 GeV y 50 GeV respectivamente.

Se han utilizado muestras de Monte carlo generadas con dos programas, Pythia and Sherpa para calcular los factores de corrección necesarios para el "unfolding" bin a bin, evaluar las incertidumbres sistemáticas y estimar las correcciones debidas al suceso subyacente y la hadronización que se han aplicado sobre los cálculos teóricos.

La sustracción del ruido de fondo se ha realizado usando una técnica orientada a datos, basada en el método de "banda lateral de 2 dimensiones", el cual depende de definir regiones de control sin sucesos de señal en el plano formado por dos variables que no están correlacionadas en el ruido de fondo. Dicho plano se ha definido usando una variable de identificación de fotones (que contiene toda la información relativa a la cascada del fotón) y la variable $E_{\mathrm{T}}^{\mathrm{iso}}$. El método utilizado tiene también en cuenta el número esperado de sucesos de la señal en las regiones de control, el cual se ha estimado utilizando muestras de Monte Carlo. La pureza como función de todos los observables estudiados es siempre superior al 90%.

Para la muestra de fotón+un jet se han medido secciones eficaces como función de E_{T}^{γ}, del momento transverso del jet con más momento transverso ($p_{\mathrm{T,lead}}^{\mathrm{jet}}$), de la masa invariante del sistema fotón-jet ($m^{\gamma j}$) y de $|\cos\theta^{\gamma j}|$. Esta última variable se corresponde con el coseno del ángulo de dispersión en el sistema de referencia del centro de masas del fotón y el jet. Las secciones eficaces han sido comparadas con cálculos teóricos al orden siguiente en teoría de perturbaciones obtenidos usando Jetphox, los cuales han sido corregidos por la influencia del suceso subyacente y de la hadronización. Las predicciones teóricas describen muy bien los datos. Se han evaluado las incertidumbres teóricas más importantes, tales como la influencia de términos superiores en el cálculo perturbativo, la influencia de la incertidumbre en la parametrización de las PDFs y la influencia de la incertidumbre en el valor de α_s.

La forma de la sección eficaz diferencial medida como funcion de $|\cos\theta^{\gamma j}|$ es consistente con la esperada en un proceso dominado por el intercambio de quarks en el canal t. Con la idea

de obtener más información sobre la dinámica del sistema fotón-jet, se ha medido también la evolución con la escala de energías de la sección eficaz diferencial como función de $|\cos\theta^{\gamma j}|$ en distintas regiones de $m^{\gamma j}$. Dichas medidas muestran un aumento del valor de la sección eficaz para $|\cos\theta^{\gamma j}| = 0.8$ conforme $m^{\gamma j}$ aumenta, el cual es descrito fielmente por los cálculos teóricos.

Para la muestras de fotón+dos jets y fotón+tres jets se ha medido la sección eficaz diferencial como función de E_{T}^{γ}, del momento transverso del segundo jet con más momento transverso ($p_{\mathrm{T,sublead}}^{\mathrm{jet}}$), del momento transverso del tercer jet con más momento transverso ($p_{\mathrm{T,subsublead}}^{\mathrm{jet}}$) y de la diferencia en el ángulo acimutal entre el fotón y los jets y los jets entre si ($\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$, $\Delta\phi^{\gamma-\mathrm{jet\,subsublead}}$, $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$, $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,subsublead}}$ and $\Delta\phi^{\mathrm{jet\,sublead}-\mathrm{jet\,subsublead}}$). Las secciones eficaces diferenciales han sido comparadas con las predicciones de Pythia y Sherpa. Las predicciones de Sherpa describen bien las secciones eficaces diferenciales, en especial, las diferencias en el ángulo acimutal en la muestra de fotón+dos jets, mientras que Pythia describe mejor las secciones eficaces como función de E_{T}^{γ} para $E_{\mathrm{T}}^{\gamma} < 900$ GeV. El hecho de que la predicción de Pythia describa peor los datos cuando hay jets adicionales tiene su explicación en la naturaleza de ambos programas de Monte Carlo; para Pythia, cuyas predicciones son $2 \to 2$, un segundo jet sólo puede provenir de la cascada partónica, mientras que para Sherpa, los elementos de matriz $2 \to 3$ y $2 \to 4$ pueden proporcionar un segundo jet directamente desde la interacción.

La evolución con la escala para las muestras de fotón+dos jets y fotón+tres jets ha sido estudiada mediante la medida de la sección eficaz diferencial como función de los ángulos acimutales entre los diversos objetos en dos regiones de E_{T}^{γ} ($130 < E_{\mathrm{T}}^{\gamma} < 300$ GeV and $E_{\mathrm{T}}^{\gamma} > 300$ GeV). Para todos los observables, la región con menor E_{T}^{γ} se comporta de forma distinta a la región sin restricciones: para $\Delta\phi^{\gamma-\mathrm{jet\,sublead}}$ y $\Delta\phi^{\gamma-\mathrm{jet\,subsublead}}$ la sección eficaz aumenta para valores de $\Delta\phi$ altos mientras que $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,sublead}}$, $\Delta\phi^{\mathrm{jet\,lead}-\mathrm{jet\,subsublead}}$ y $\Delta\phi^{\mathrm{jet\,sublead}-\mathrm{jet\,subsublead}}$ se ven suprimidas.

Se ha medido por primera vez en ATLAS la coherencia de color, en la cual, el flujo de color entre partones en el estado inicial o final modifica el patrón de radiación partónica alrededor de partones coloreados. Con ese propósito se han definido dos sub-muestras de sucesos con al menos dos jets en el estado final en las cuales el segundo jet con más momento transverso se encuentra cerca del fotón o cerca del jet con más momento transverso. Se han medido las secciones eficaces diferenciales como función del ángulo en el plano $\eta - \phi$ entre el segundo jet con más momento transverso y el fotón cuando el primero se encuentra cerca del fotón (β^{γ}) y como función del ángulo en el plano $\eta - \phi$ entre el segundo jet con más momento transverso y el jet con más momento transverso cuando el primero se encuentra cerca del jet con más momento transverso ($\beta_{\mathrm{lead}}^{\mathrm{jet}}$). Dichas secciones eficaces proporcionan una medida de la dirección de la radiación alrededor de ambos objetos. Ambas secciones eficaces se comportan de forma muy distinta: $d\sigma/d\beta_{\mathrm{lead}}^{\mathrm{jet}}$ aumenta conforme $\beta_{\mathrm{lead}}^{\mathrm{jet}}$ aumenta, lentamente para $\beta_{\mathrm{lead}}^{\mathrm{jet}} < \pi/2$ y más rápido para $\beta_{\mathrm{lead}}^{\mathrm{jet}} > \pi/2$, mientras que $d\sigma/d\beta^{\gamma}$ presenta un máximo en $\beta^{\gamma} \approx \pi/2$. Las predicciones de Sherpa describen bien las secciones eficaces medidas. Con la intención de confirmar la presencia de coherencia de color utilizando solamente datos, se ha medido el cociente entre las dos secciones diferenciales anteriormente descritas; dicho cociente muestra de forma clara un aumento de la sección eficaz como función de $\beta_{\mathrm{lead}}^{\mathrm{jet}}$ en $\beta_{\mathrm{lead}}^{\mathrm{jet}} = 0$ y en $\beta_{\mathrm{lead}}^{\mathrm{jet}} = \pi$ con respecto a la sección eficaz como función de β^{γ}. Este comportamiento indica un aumento de la radiación en la dirección de los partones del estado inicial alrededor del jet con momento transverso más alto con respecto a una partícula incolora (el fotón), el cual es consistente con el esperado debido a la coherencia de color.

En resumen, se han presentado dos análisis que mejoran el conocimiento de las partículas fundamentales y de las interacciones en la frontera de energía proporcionada por el LHC. Las medidas de secciones eficaces para sucesos de fotón en asociación con jets y su comparación con las predicciones teóricas proporcionan un test muy riguroso de la cromodinámica cuántica

y permiten obtener nueva información para caracterizar una de las fuentes de ruido de fondo más importantes para la medida de producción del bosón de Higgs, en el canal en el cual éste se desintegra en dos fotones, y de cualquier otra hipotética partícula que se desintegre en fotones. Asimismo, el método desarrollado para la identificación de partículas con un alto momento transverso proporciona una técnica eficiente y fidedigna para las búsquedas de nuevas partículas muy masivas de una manera que no presenta singularidades en los límites infrarrojo y colineal.

Bibliography

[1] S.L. Glashow. "Partial Symmetries of Weak Interactions".
In: *Nucl. Phys.* 22 (1961), pp. 579–588. DOI: 10.1016/0029-5582(61)90469-2.

[2] S. Weinberg. "A Model of Leptons". In: *Phys. Rev. Lett.* 19 (1967), pp. 1264–1266.
DOI: 10.1103/PhysRevLett.19.1264.

[3] A. Salam. *In Elementary particle physics: relativistic groups and analyticity.*
ed. N, Svartholm, Almquist and Wiksell, 1968.

[4] P. W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons".
In: *Phys. Rev. Lett.* 13 (16 Oct. 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508.
URL: http://link.aps.org/doi/10.1103/PhysRevLett.13.508.

[5] P. W. Higgs. "Broken symmetries, massless particles and gauge fields".
In: *Phys. Lett.* 12 (1964), pp. 132–133. DOI: 10.1016/0031-9163(64)91136-9.

[6] P. W. Higgs. "Spontaneous Symmetry Breakdown without Massless Bosons".
In: *Phys. Rev.* 145 (4 May 1966), pp. 1156–1163. DOI: 10.1103/PhysRev.145.1156.
URL: http://link.aps.org/doi/10.1103/PhysRev.145.1156.

[7] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons".
In: *Phys. Rev. Lett.* 13 (9 Aug. 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321.
URL: http://link.aps.org/doi/10.1103/PhysRevLett.13.321.

[8] ATLAS Collaboration. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC".
In: *Phys. Lett.* B716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020.
arXiv: 1207.7214 [hep-ex].

[9] ATLAS Collaboration. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Phys. Lett.* B716 (2012), pp. 30–61.
DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[10] *The Nobel Prize in Physics 2013.* The Royal Swedish Academy of Sciences, Oct. 2013.
URL: http://www.nobelprize.org/nobel_prizes/physics/laureates/2013/.

[11] H.D. Politzer. "Reliable perturbative results for strong interactions".
In: *Phys. Rev. Lett.* 30 (1973), pp. 1346–1349. DOI: 10.1103/PhysRevLett.30.1346.

[12] D.J. Gross and F. Wilczek. "Ultraviolet behavior of non-abelian gauge theories".
In: *Phys. Rev. Lett.* 30 (1973), pp. 1343–1346. DOI: 10.1103/PhysRevLett.30.1343.

[13] H. Fritzsch, M. Gell-Mann, and H. Leutwyler.
"Advantages of the Color Octet Gluon Picture". In: *Phys. Lett.* B47 (1973), pp. 365–368.
DOI: 10.1016/0370-2693(73)90625-4.

[14] J. C. Collins and D. E. Soper. "Parton Distribution and Decay Functions".
In: *Nucl. Phys.* B194 (1982), p. 445. DOI: 10.1016/0550-3213(82)90021-9.

[15] J. C. Collins, D. E. Soper, and G. F. Sterman.
"Transverse Momentum Distribution in Drell-Yan Pair and W and Z Boson Production".
In: *Nucl. Phys.* B250 (1985), p. 199. DOI: 10.1016/0550-3213(85)90479-1.

[16] G. Altarelli and G. Parisi. "Asymptotic Freedom in Parton Language". In: *Nucl. Phys.* B126 (1977), p. 298. DOI: 10.1016/0550-3213(77)90384-4.

[17] V. N. Gribov and L. N. Lipatov. "Deep inelastic e p scattering in perturbation theory". In: *Sov. J. Nucl. Phys.* 15 (1972). [Yad. Fiz.15,781(1972)], pp. 438–450.

[18] Y. L. Dokshitzer. "Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics." In: *Sov. Phys. JETP* 46 (1977). [Zh. Eksp. Teor. Fiz.73,1216(1977)], pp. 641–653.

[19] A. Buckley and M. Whalley. "HepData reloaded: Reinventing the HEP data archive". In: *PoS* ACAT2010 (2010), p. 067. arXiv: 1006.0517 [hep-ex].

[20] P. M. Nadolsky et al. "Implications of CTEQ global analysis for collider observables". In: *Phys. Rev.* D78 (2008), p. 013004. DOI: 10.1103/PhysRevD.78.013004. arXiv: 0802.0007 [hep-ph].

[21] K. A. Olive et al. "Review of Particle Physics". In: *Chin. Phys.* C38 (2014), p. 090001. DOI: 10.1088/1674-1137/38/9/090001.

[22] *The Nobel Prize in Physics 2004*. The Royal Swedish Academy of Sciences, Oct. 2004. URL: http://www.nobelprize.org/nobel_prizes/physics/laureates/2004/.

[23] M. S. S. Catani Y.L. Dokshitzer and B.R. Webber. "Longitudinally invariant K_t clustering algorithms for hadron hadron collisions". In: *Nucl. Phys.* B406 (1993), pp. 187–224. DOI: 10.1016/0550-3213(93)90166-M.

[24] G. S. M. Cacciari and G. Soyez. "The anti-k_t jet clustering algorithm". In: *JHEP* 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph].

[25] M. Kobayashi and T. Maskawa. "CP Violation in the Renormalizable Theory of Weak Interaction". In: *Prog. Theor. Phys.* 49 (1973), pp. 652–657. DOI: 10.1143/PTP.49.652.

[26] F. Abe et al. "Observation of top quark production in $\bar{p}p$ collisions". In: *Phys. Rev. Lett.* 74 (1995), pp. 2626–2631. DOI: 10.1103/PhysRevLett.74.2626. arXiv: hep-ex/9503002.

[27] ATLAS and CMS and CDF and D0 Collaborations. "First combination of Tevatron and LHC measurements of the top-quark mass". In: (2014). arXiv: 1403.4427 [hep-ex].

[28] M. Cacciari et al. "Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation". In: *Phys.Lett.* B710 (2012), pp. 612–622. DOI: 10.1016/j.physletb.2012.03.013. arXiv: 1111.5869 [hep-ph].

[29] M. Beneke et al. "Hadronic top-quark pair production with NNLL threshold resummation". In: *Nucl.Phys.* B855 (2012), pp. 695–741. DOI: 10.1016/j.nuclphysb.2011.10.021. arXiv: 1109.1536 [hep-ph].

[30] P. Baernreuther, M. Czakon, and A. Mitov. "Percent Level Precision Physics at the Tevatron: First Genuine NNLO QCD Corrections to $q\bar{q} \to t\bar{t} + X$". In: *Phys.Rev.Lett.* 109 (2012), p. 132001. DOI: 10.1103/PhysRevLett.109.132001. arXiv: 1204.5201 [hep-ph].

[31] M. Czakon and A. Mitov. "Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders". In: (2011). arXiv: 1112.5675 [hep-ph].

[32] A. Martin et al. "Parton distributions for the LHC". In: *Eur.Phys.J.* C63 (2009), pp. 189–285. DOI: 10.1140/epjc/s10052-009-1072-5. arXiv: 0901.0002 [hep-ph].

[33] N. Cabibbo. "Unitary Symmetry and Leptonic Decays".
In: *Phys. Rev. Lett.* 10 (12 June 1963), pp. 531–533.
DOI: 10.1103/PhysRevLett.10.531.
URL: http://link.aps.org/doi/10.1103/PhysRevLett.10.531.

[34] Z. Belghobsi et al. "Photon-jet correlations and constraints on fragmentation functions".
In: *Phys. Rev. D* 79 (11 June 2009), p. 114024. DOI: 10.1103/PhysRevD.79.114024.
URL: http://link.aps.org/doi/10.1103/PhysRevD.79.114024.

[35] S. Catani et al. "Cross-section of isolated prompt photons in hadron hadron collisions".
In: *JHEP* 05 (2002), p. 028. DOI: 10.1088/1126-6708/2002/05/028.
arXiv: hep-ph/0204023 [hep-ph].

[36] F. Abe et al. "Evidence for color coherence in pp collisions at $\sqrt{s} = 1.8$ TeV".
In: *Phys. Rev. D* 50 (9 Nov. 1994), pp. 5562–5579. DOI: 10.1103/PhysRevD.50.5562.
URL: http://link.aps.org/doi/10.1103/PhysRevD.50.5562.

[37] B. Abbott et al.
"Color coherent radiation in multijet events from $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV".
In: *Phys. Lett.* B414 (1997), pp. 419–427. DOI: 10.1016/S0370-2693(97)01190-8.
arXiv: hep-ex/9706012 [hep-ex].

[38] M. Akrawy et al. "A study of coherence of soft gluons in hadron jets".
In: *Physics Letters B* 247.4 (1990), pp. 617–628. ISSN: 0370-2693.
DOI: http://dx.doi.org/10.1016/0370-2693(90)91911-T.
URL: http://www.sciencedirect.com/science/article/pii/037026939091911T.

[39] Y. L. Dokshitzer et al. "QCD coherence in high-energy reactions".
In: *Rev. Mod. Phys.* 60 (2 Apr. 1988), pp. 373–388. DOI: 10.1103/RevModPhys.60.373.
URL: http://link.aps.org/doi/10.1103/RevModPhys.60.373.

[40] O. S. Brüning et al. *LHC Design Report*. Geneva: CERN, 2004.
URL: http://cds.cern.ch/record/782076.

[41] L. Evans. "The Large Hadron Collider". In: *New J. Phys.* 9.9 (2007), p. 335.
URL: http://stacks.iop.org/1367-2630/9/i=9/a=335.

[42] ATLAS Collaboration. *ATLAS detector and physics performance*.
Technical Design Report ATLAS. Geneva: CERN, 1999.
URL: http://cds.cern.ch/record/391176.

[43] ATLAS Collaboration. "The ATLAS Experiment at the CERN Large Hadron Collider".
In: *JINST* 3.08 (2008), S08003.
URL: http://stacks.iop.org/1748-0221/3/i=08/a=S08003.

[44] T. Sjostrand, S. Mrenna, and P. Z. Skands. "PYTHIA 6.4 Physics and Manual".
In: *JHEP* 0605 (2006), p. 026. DOI: 10.1088/1126-6708/2006/05/026.
arXiv: hep-ph/0603175 [hep-ph].

[45] T. Gleisberg et al. "Event generation with SHERPA 1.1". In: *JHEP* 02 (2009), p. 007.
DOI: 10.1088/1126-6708/2009/02/007. arXiv: 0811.4622 [hep-ph].

[46] M. L. Mangano et al.
"ALPGEN, a generator for hard multiparton processes in hadronic collisions".
In: *JHEP* 0307 (2003), p. 001. DOI: 10.1088/1126-6708/2003/07/001.
arXiv: hep-ph/0206293 [hep-ph].

[47] S. Frixione and B. R. Webber.
"Matching NLO QCD computations and parton shower simulations".
In: *JHEP* 0206 (2002), p. 029. DOI: 10.1088/1126-6708/2002/06/029.
arXiv: hep-ph/0204244 [hep-ph].

[48] B. P. Kersevan and E. Richter-Was. "The Monte Carlo event generator AcerMC versions 2.0 to 3.8 with interfaces to PYTHIA 6.4, HERWIG 6.5 and ARIADNE 4.1". In: *Comput.Phys.Commun.* 184 (2013), pp. 919–985. DOI: 10.1016/j.cpc.2012.10.032. arXiv: hep-ph/0405247 [hep-ph].

[49] G. Corcella et al. "HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)". In: *JHEP* 0101 (2001), p. 010. DOI: 10.1088/1126-6708/2001/01/010. arXiv: hep-ph/0011363 [hep-ph].

[50] B. Andersson et al. "Parton Fragmentation and String Dynamics". In: *Phys. Rept.* 97 (1983), pp. 31–145. DOI: 10.1016/0370-1573(83)90080-7.

[51] B.R. Webber. "A QCD Model for Jet Fragmentation Including Soft Gluon Interference". In: *Nucl. Phys.* B238 (1984), p. 492. DOI: 10.1016/0550-3213(84)90333-X.

[52] T. Sjostrand and M. van Zijl. "A Multiple Interaction Model for the Event Structure in Hadron Collisions". In: *Phys. Rev.* D36 (1987), p. 2019. DOI: 10.1103/PhysRevD.36.2019.

[53] J. Butterworth, J. R. Forshaw, and M. Seymour. "Multiparton interactions in photoproduction at HERA". In: *Z.Phys.* C72 (1996), pp. 637–646. DOI: 10.1007/s002880050286. arXiv: hep-ph/9601371 [hep-ph].

[54] S. Agostinelli et al. "GEANT4: A Simulation toolkit". In: *Nucl.Instrum.Meth.* A506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.

[55] H.-L. Lai et al. "New parton distributions for collider physics". In: *Phys. Rev.* D 82 (2010), p. 074024. arXiv: 1007.2241 [hep-ph].

[56] ATLAS Collaboration. *New ATLAS event generator tunes to 2010 data.* Tech. rep. ATL-PHYS-PUB-2011-008. Geneva: CERN, Apr. 2011. URL: http://cds.cern.ch/record/1345343.

[57] J. Pumplin et al. "New generation of parton distributions with uncertainties from global QCD analysis". In: *JHEP* 0207 (2002), p. 012. arXiv: 0201195 [hep-ph].

[58] A. Sherstnev and R. S. Thorne. "Different PDF approximations useful for LO Monte Carlo generators". In: *Proceedings, 16th International Workshop on Deep Inelastic Scattering and Related Subjects (DIS 2008)*. 2008, p. 149. DOI: 10.3360/dis.2008.149. arXiv: 0807.2132 [hep-ph]. URL: http://inspirehep.net/record/790666/files/arXiv:0807.2132.pdf.

[59] ATLAS Collaboration. *ATLAS tunes of PYTHIA 6 and Pythia 8 for MC11.* Tech. rep. ATL-PHYS-PUB-2011-009. Geneva: CERN, July 2011. URL: http://cds.cern.ch/record/1363300.

[60] T. Sjostrand, S. Mrenna, and P. Z. Skands. "A Brief Introduction to PYTHIA 8.1". In: *Comput. Phys. Commun.* 178 (2008), pp. 852–867. DOI: 10.1016/j.cpc.2008.01.036. arXiv: 0710.3820 [hep-ph].

[61] ATLAS Collaboration. *Summary of ATLAS Pythia 8 tunes.* Tech. rep. ATL-PHYS-PUB-2012-003. Geneva: CERN, Aug. 2012. URL: http://cds.cern.ch/record/1474107.

[62] ATLAS Collaboration. *Expected performance of the ATLAS experiment: detector, trigger and physics.* Geneva: CERN, 2009. URL: http://cds.cern.ch/record/1125884.

[63] ATLAS Collaboration.
"Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data".
In: *Eur. Phys. J.* C74 (2014), p. 2941. DOI: 10.1140/epjc/s10052-014-2941-0.
arXiv: 1404.2240 [hep-ex].

[64] ATLAS Collaboration. *Expected photon performance in the ATLAS experiment.*
Tech. rep. ATL-PHYS-PUB-2011-007. Geneva: CERN, Apr. 2011.
URL: https://cds.cern.ch/record/1345329.

[65] ATLAS Collaboration.
"Monitoring and data quality assessment of the ATLAS liquid argon calorimeter".
In: *JINST* 9 (2014), P07024. DOI: 10.1088/1748-0221/9/07/P07024.
arXiv: 1405.3768 [hep-ex].

[66] ATLAS Collaboration. *Expected electron performance in the ATLAS experiment.*
Tech. rep. ATL-PHYS-PUB-2011-006. Geneva: CERN, Apr. 2011.
URL: http://cds.cern.ch/record/1345327.

[67] D. Banfi, L. Carminati, and L. Mandelli.
Calibration of the ATLAS electromagnetic calorimeter using calibration hits.
Tech. rep. ATL-LARG-PUB-2007-012. ATL-COM-LARG-2007-007.
Geneva: CERN, July 2007. URL: http://cds.cern.ch/record/1046248.

[68] ATLAS Collaboration. *Electron and photon reconstruction and identification in ATLAS: expected performance at high energy and results at 900 GeV.*
Tech. rep. ATLAS-CONF-2010-005. Geneva: CERN, June 2010.
URL: https://cds.cern.ch/record/1273197.

[69] ATLAS Collaboration. "Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data". In: *Eur. Phys. J.* C72 (2012), p. 1909.
DOI: 10.1140/epjc/s10052-012-1909-1. arXiv: 1110.3174 [hep-ex].

[70] ATLAS Collaboration. "Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data". In: *Eur. Phys. J.* C74.10 (2014), p. 3071.
DOI: 10.1140/epjc/s10052-014-3071-4. arXiv: 1407.5063 [hep-ex].

[71] M. Cacciari, G. P. Salam, and S. Sapeta.
"On the characterisation of the underlying event". In: *JHEP* 04 (2010), p. 065.
DOI: 10.1007/JHEP04(2010)065. arXiv: 0912.4926 [hep-ph].

[72] M. Cacciari, G. P. Salam, and G. Soyez. "FastJet User Manual".
In: *Eur. Phys. J.* C72 (2012), p. 1896. DOI: 10.1140/epjc/s10052-012-1896-2.
arXiv: 1111.6097 [hep-ph].

[73] W. Lampl et al. *Calorimeter Clustering Algorithms: Description and Performance.*
Tech. rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003.
Geneva: CERN, Apr. 2008. URL: http://cds.cern.ch/record/1099735.

[74] ATLAS Collaboration. "Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector".
In: *Eur. Phys. J.* C75 (2015), p. 17. DOI: 10.1140/epjc/s10052-014-3190-y.
arXiv: 1406.0076 [hep-ex].

[75] M. Aharrouche et al. "Measurement of the response of the ATLAS liquid argon barrel calorimeter to electrons at the 2004 combined test-beam".
In: *Nucl. Instrum. Meth.* A614 (2010), pp. 400–432. DOI: 10.1016/j.nima.2009.12.055.

[76] ATLAS Collaboration. "Characterisation and mitigation of beam-induced backgrounds observed in the ATLAS detector during the 2011 proton-proton run".
In: *JINST* 8 (2013), P07004. DOI: 10.1088/1748-0221/8/07/P07004.
arXiv: 1303.0223 [hep-ex].

[77] ATLAS Collaboration. *Monte Carlo Calibration and Combination of In-situ Measurements of Jet Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS.* Tech. rep. ATLAS-CONF-2015-037. Geneva: CERN, Aug. 2015. URL: http://cds.cern.ch/record/2044941.

[78] ATLAS Collaboration. "Luminosity Determination in *pp* Collisions at $\sqrt{s} = 7$ TeV Using the ATLAS Detector at the LHC". In: *Eur. Phys. J.* C71 (2011), p. 1630. DOI: 10.1140/epjc/s10052-011-1630-5. arXiv: 1101.2185 [hep-ex].

[79] M. Cacciari and G. P. Salam. "Pileup subtraction using jet areas". In: *Phys. Lett.* B659 (2008), pp. 119–126. DOI: 10.1016/j.physletb.2007.09.077. arXiv: 0707.1378 [hep-ph].

[80] ATLAS Collaboration. *Pile-up subtraction and suppression for jets in ATLAS.* Tech. rep. ATLAS-CONF-2013-083. Geneva: CERN, Aug. 2013. URL: https://cds.cern.ch/record/1570994.

[81] ATLAS Collaboration. *Jet global sequential corrections with the ATLAS detector in proton-proton collisions at sqrt(s) = 8 TeV.* Tech. rep. ATLAS-CONF-2015-002. Geneva: CERN, Mar. 2015. URL: https://cds.cern.ch/record/2001682.

[82] ATLAS Collaboration. *Data-driven determination of the energy scale and resolution of jets reconstructed in the ATLAS calorimeters using dijet and multijet events at* $\sqrt{s} = 8$ *TeV.* Tech. rep. ATLAS-CONF-2015-017. Geneva: CERN, Apr. 2015. URL: https://cds.cern.ch/record/2008678.

[83] ATLAS Collaboration. *Measurement of the b-tag Efficiency in a Sample of Jets Containing Muons with 5 fb-1 of Data from the ATLAS Detector.* Tech. rep. ATLAS-CONF-2012-043. Geneva: CERN, Mar. 2012. URL: http://cds.cern.ch/record/1435197.

[84] ATLAS Collaboration. *Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data.* Tech. rep. ATLAS-CONF-2011-102. Geneva: CERN, July 2011. URL: http://cds.cern.ch/record/1369219.

[85] G. Piacquadio and C. Weiser. "A new inclusive secondary vertex algorithm for b-jet tagging in ATLAS". In: *Journal of Physics: Conference Series* 119.3 (2008), p. 032032. URL: http://stacks.iop.org/1742-6596/119/i=3/a=032032.

[86] ATLAS Collaboration. "Performance of jet substructure techniques for large-R jets in proton-proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector". In: *JHEP* 09 (2013), p. 076. DOI: 10.1007/JHEP09(2013)076. arXiv: 1306.4945 [hep-ex].

[87] ATLAS Collaboration. *Performance of boosted top quark identification in 2012 ATLAS data.* Tech. rep. ATLAS-CONF-2013-084. Geneva: CERN, Aug. 2013. URL: http://cds.cern.ch/record/1571040.

[88] ATLAS Collaboration. "Improved luminosity determination in *pp* collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC". In: *Eur. Phys. J.* C 73 (2013), p. 2518. arXiv: 1302.4393 [hep-ex].

[89] ATLAS Collaboration. "Measurement of the inclusive isolated prompt photon cross section in *pp* collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector". In: *Phys. Rev.* D 83 (2011), p. 052005. arXiv: 1012.4389 [hep-ex].

[90] ATLAS Collaboration. "Dynamics of isolated-photon plus jet production in *pp* collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector". In: *Nucl. Phys.* B 875 (2013), 483 [arXiv:1307.6795].

[91] M. Stoebe et al. *Measurement of the inclusive isolated prompt photon cross section in pp collisions at* $\sqrt{s} = 8$ *TeV with the ATLAS detector using 20.2* fb^{-1}.
ATL-COM-PHYS-2014-744.

[92] S. Catani et al. "Cross section of isolated prompt photons in hadron-hadron collisions". In: *JHEP* 0205 (2002), p. 028. arXiv: 0204023 [hep-ph].

[93] P. Aurenche et al. "A New critical study of photon production in hadronic collisions". In: *Phys. Rev.* D 73 (2006), p. 094007. arXiv: 0602133 [hep-ph].

[94] L. Bourhis, M. Fontannaz and J.Ph. Guillet.
"Quark and gluon fragmentation functions into photons".
In: *Eur. Phys. J.* C 2 (1998), p. 529. arXiv: 9704447 [hep-ph].

[95] A.D. Martin and others. "Uncertainties on alpha(S) in global PDF analyses and implications for predicted hadronic cross sections". In: *Eur. Phys. J.* C 64 (2009), p. 653. arXiv: 0905.3531 [hep-ph].

[96] G. D'Agostini. "A Multidimensional unfolding method based on Bayes' theorem".
In: *Nucl. Instrum. Meth.* A362 (1995), pp. 487–498.
DOI: 10.1016/0168-9002(95)00274-X.

[97] T. Adye. "Unfolding algorithms and tests using RooUnfold". In: arXiv:1105.1160 (May 2011). Comments: 6 pages, 5 figures, presented at PHYSTAT 2011, CERN, Geneva, Switzerland, January 2011, to be published in a CERN Yellow Report, 6 p.
URL: https://cds.cern.ch/record/1349242.

Agradecimientos

Llegado el final de esta etapa, no puedo dejar de dar las gracias a toda la gente con la que he compartido camino estos últimos años:

A mis directores de tesis, Juan y Claudia. Gracias por todas las lecciones, consejos y ayuda sin las cuales, evidentemente, no habría tesis. Gracias también por la infinita paciencia ante mi evolución de tortuga, por haberme enseñado a diferenciar entre lo que importa y lo que no importa a la hora de hacer física y por vuestra cercanía y amistad.

A José del Peso y Edu. Gracias por guiarme en mis primeros pasos, siempre complicados, en el CERN.

To all the great people I met in the LAr operations team: Hass, Claire Lee, Stephanie, Nikiforos, Pavol, Bertrand, Sergey, Isabelle, Emmanuel, Luis and many others. Thanks for teaching me, trusting me and making all the long nights and early mornings worth it.

A toda la gente con la que he compartido comidas, cafés y buenos ratos en Ginebra: Joel, Luis, Neus, Javi, Tito, Josu, Dani, María, Sebas, Jordi y muchos más. Gracias porque aunque sea propenso a desaparecer, siempre me he sentido aceptado cuando he vuelto. Gracias especialmente a Sandra, mi apoyo friki en el exilio.

A Laura y Roberto. Gracias por haber conseguido que venir a la UAM sea volver a casa, aunque nos veamos una vez al año.

A mis padres. Gracias por haberme dejado siempre elegir mi propio camino sin ponerme pegas. Gracias por un millón de cosas que no se pueden resumir en este texto. Soy quien soy gracias a vosotros.

A mi abuela. Aunque no vaya a mencionarte en los agradecimientos del premio Nobel, al menos te menciono en los de la tesis. Gracias por no haberme echado nunca en cara que sólo tengas nieto de pascuas a ramos.

A todos los que se preocupan por mí y me cuidan siempre que pueden, sean familia o no, de Ginebra y de Madrid: Begoña, Sandra, Alicia, Mariano, Teresa, Isabel, José Miguel, Roberto y tantos otros.

A mis amigos: Jorge, Gerardo y Jesús. Lo que me habéis ofrecido vosotros, toda la amistad, el apoyo incondicional y los buenos momentos no tienen precio. Me vale y me sobra con saber que siempre estaréis ahí. Gracias por ser la constante esté donde esté y ocurra lo que ocurra.

A Victoria. Gracias por todo y más, por haberme seguido a la aventura y por aguantar todas mis cosas sin juzgarme. Somos como somos y lo que tenemos no lo cambio por nada.

A Chazanne y Cone.